

SYSTEMATISCHE FEHLER IN STICHPROBENERHEBUNGEN UND IHRE AUSWIRKUNGEN AUF SCHÄTZAUSSAGEN

DISSERTATION
zur Erlangung des Grades
eines Doktors der
wirtschaftlichen Staatswissenschaften
(Dr. rer. pol.)
des Fachbereichs Rechts- und
Wirtschaftswissenschaften
der
Johannes Gutenberg-Universität Mainz

vorgelegt von
Diplom-Handelslehrer
Johann Strauch

1984

SYSTEMATISCHE FEHLER IN
STICHPROBENZENTRIFERUNGEN
UND IHRE AUSWIRKUNGEN
AUF SCHATZAUSSAGEN



Digitized by the Internet Archive
in 2026

SYSTEMATISCHE FEHLER IN STICHPROBENERHEBUNGEN UND IHRE AUSWIRKUNGEN AUF SCHÄTZAUSSAGEN

DISSERTATION
zur Erlangung des Grades
eines Doktors der
wirtschaftlichen Staatswissenschaften
(Dr. rer. pol.)
des Fachbereichs Rechts- und
Wirtschaftswissenschaften
der
Johannes Gutenberg-Universität Mainz

vorgelegt von
Diplom-Handelslehrer
Johann Strauch

1984

Erstberichterstatter: Prof. Dr. Erich Härtter

Datum der Berichterstattung: 31. 10. 1984

Zweitberichterstatter: Prof. Dr. Peter M. Schulze

Datum der Berichterstattung: 19. 11. 1984

Tag der mündlichen Prüfung: 19. 12. 1984

Inhaltsverzeichnis

	Seite
1. Einleitung	1
2. Fehler in statistischen Untersuchungen	5
2.1 Fehlerquellen in statistischen Untersuchungen	5
2.1.1 Überblick	5
2.1.2 Stichprobenbedingte Fehlerquellen	9
2.1.3 Nicht-stichprobenbedingte Fehlerquellen	11
2.1.3.1 Fehlerquellen bei der Erfassung der Untersuchungseinheiten	11
2.1.3.2 Fehlerquellen bei der inhaltlichen Erfassung der Merkmalsausprägungen	12
2.2 Begriffliche Abgrenzung verschiedener Fehlerarten	19
2.2.1 Zufallsfehler, systematischer Fehler und Verzerrung	19
2.2.2 Die Stichprobenauswahl unter Berücksichtigung des Antwortverhaltens der Untersuchungseinheiten	23
2.2.2.1 Die Zufallsstichprobe bei korrekten individuellen Antworten	23
2.2.2.2 Die Antworten der Untersuchungseinheiten als Zufallsvariable	26
2.2.3 Der Einfluß des Antwortverhaltens auf den Zufallsfehler und die Verzerrung der Schätzung	31
2.2.4 Tatsächlicher Fehler und Mean Square Error	34
2.3 Eigenschaften der Fehler in statistischen Untersuchungen	39
2.3.1 Eigenschaften des Zufallsfehlers einer Schätzung	39
2.3.2 Eigenschaften der Verzerrung einer Schätzung	42

	Seite
2.4 Kontrollmaßnahmen und ihre Bedeutung für Stichprobenerhebungen	46
3. Auswirkungen systematischer Fehler auf Schätzungen bei korrekten indi- viduellen Antworten	52
3.1 Grundlagen	52
3.2 Die Verzerrung bei einfacher Zufalls- auswahl	54
3.2.1 Das Modell der einfachen Zufallsauswahl	54
3.2.1.1 Die Stichprobe als Zufallsvektor	54
3.2.1.2 Der Mittelwertschätzer bei einfacher Zufallsauswahl	58
3.2.1.3 Der Zufallsfehler des Schätzers	61
3.2.2 Die Beschreibung systematischer Fehler im Modell der einfachen Zufallsauswahl	62
3.2.2.1 Die Entstehung der Verzerrung	62
3.2.2.2 Fehlerhafte Auswahlgrundlage	63
3.2.2.3 Systematische Fehler im Auswahlprozeß	65
3.2.3 Der tatsächliche Fehler bei einfacher Zufallsauswahl	66
3.2.4 Auswirkungen der systematischen Fehler auf die Schätzaussage bei einfacher Zufallsauswahl	69
3.2.4.1 Die Konfidenzaussage bei verzerrter Schätzung	69
3.2.4.2 Ein Versuch, die Auswirkungen einer Verzerrung über den Stichprobenumfang zu verdeutlichen	74
3.2.5 Möglichkeiten einer Berücksichtigung der Verzerrung in der Schätzaussage	80
3.3 Die Verzerrung beim Schichtenauswahl- verfahren	85
3.3.1 Das Schichtenauswahlverfahren	85
3.3.1.1 Grundlegende Notationen	85
3.3.1.2 Die Mittelwertschätzung bei geschichte- ter Auswahl	88
3.3.2 Die Beschreibung systematischer Fehler im Schichtenauswahlverfahren	92
3.3.3 Auswirkungen der Verzerrung auf die Schätzaussage beim Schichtenauswahl- verfahren	94

	Seite
3.3.3.1 Das Konfidenzintervall bei verzerrtem geschichtetem Schätzer	94
3.3.3.2 Ein Vergleich zwischen verzerrter und unverzerrter geschichteter Schätzung über den Stichprobenumfang	96
3.4 Vergleich zwischen einfacher Zufallsauswahl und Schichtenauswahl	97
3.4.1 Der Schichteneffekt	97
3.4.2 Ein Vergleich zwischen einfacher Zufallsauswahl und Schichtenauswahl bei verzerrten Schätzungen	101
3.4.2.1 Die Verzerrung bei einfacher und geschichteter Zufallsauswahl	101
3.4.2.2 Ein quantitativer Vergleich der Auswirkungen der Verzerrung auf die Schätzaussage bei einfacher und geschichteter Zufallsauswahl	104
4. Die Mittelwertschätzung unter Berücksichtigung systematischer Fehler in den Antworten	113
4.1 Abgrenzungen	113
4.2 Die einfache Zufallsauswahl bei Antwortfehlern	119
4.2.1 Die Stichprobe bei Antwortfehlern	119
4.2.2 Die Mittelwertschätzung bei Antwortfehlern	123
4.2.2.1 Erwartungswert und Varianz des Mittelwertschätzers bei Antwortfehlern	123
4.2.2.2 Die Abschätzbarkeit des Zufallsfehlers bei Antwortfehlern	131
4.2.3 Auswirkungen der Antwortfehler auf die Schätzaussage bei einfacher Zufallsauswahl	134
4.2.4 Eigenschaften des Zufallsfehlers der Schätzung bei Antwortfehlern	137
4.3 Antwortfehler im Schichtenauswahlverfahren	140
4.3.1 Grundlegende Notationen	140
4.3.2 Die geschichtete Mittelwertschätzung bei Antwortfehlern	142

	Seite
4.3.2.1 Erwartungswert und Varianz des geschichteten Mittelwertschätzers bei Antwortfehlern	142
4.3.2.2 Die Aufteilung der Stichprobe auf die Schichten unter Berücksichtigung der Antwortfehler	146
4.3.3 Auswirkungen der Antwortfehler auf die Schätzaussage beim Schichten- auswahlverfahren	147
4.4 Vergleich zwischen einfacher und geschichteter Auswahl unter Berücksichtigung systematischer Fehler in den Antworten	150
4.4.1 Annahmen, die dem Vergleich zugrunde liegen	150
4.4.2 Die Beziehung zwischen der Varianz in der Gesamtheit und den Varianzen innerhalb der Schichten	152
4.4.3 Der Schichteneffekt unter Berücksichtigung der Antwortfehler	155
5. Schlußbetrachtungen	159

Literaturverzeichnis

- Anderson, O.: Über den Umgang mit systematischen Fehlern, in : Statistische Vierteljahresschrift, Vol. 7, 1954, S. 38 - 44.
- Anderson, O.: Probleme der Statistischen Methodenlehre in den Sozialwissenschaften, 5. Aufl., Würzburg 1965.
- Atteslander, P.: Methoden der empirischen Sozialforschung, 4. Aufl., Berlin, New York 1975.
- Bley Müller, J., Gehlert, G., Gülicher, H.: Statistik für Wirtschaftswissenschaftler, 3. Aufl., München 1983.
- Cochran, W. G.: Stichprobenverfahren, Berlin, New York 1972.
- Dalenius, T. u. Hodges, J.: The Choice of Stratification Points, in: Skandinavisk Aktuarietidskrift, 1957, S. 198 - 203.
- Deming, W.: Some Theory of Sampling, New York, London 1950.
- Esenwein-Rothe, J.: Ungenauigkeiten und Fehler in wirtschaftsstatistischen Daten, in: Statistische Hefte, 11. Jahrg. 1970, S. 160 - 193.
- Hansen, M. H., Hurwitz, W. N., Madow, W. G.: Sample Survey Methods and Theory, Vol. I u. II, New York, London, Sydney 1953.
- Hansen, M. H., Hurwitz, W. N., Marks, E. S., Muldin, P. W.: Response Errors in Survey, in: Journal of the American Statistical Association, Vol. 46, 1951, S. 147 - 190.
- Holm, K.: Das Modell des Untersuchungsgegenstandes in: Die Befragung, hrsg. v. K. Holm, Bd. 1, München 1975 a, S. 14 - 21.
- Holm K.: Die Frage, in: Die Befragung, hrsg. v. K. Holm, Bd. 1, München 1975 b, S. 32 - 91.

- Kallmeyer, H.: Über Fehler, Fehlerausgleich und Fehlerfortpflanzung in der Sozialstatistik, in: Allgemeines statistisches Archiv, Vol. 40, 1956, S. 19 - 37.
- v. Kirschhofer-Bozenhardt, A., Kaplitza, G.: Der Fragebogen, in: Die Befragung, Bd. 1, hrsg. v. K. Holm. München 1975, S. 92 - 126.
- Koop, I. C.: Notes for a Unified Theory of Estimation for Sample Surveys Taking into Account Response Errors, in: Metrika, Vol. 21, 1974, S. 19 - 39.
- Krug, W., Nourney, M.: Wirtschafts- und Sozialstatistik: Gewinnung von Daten, München, Wien 1982
- Mahalanobis, P. C.: Recent experiments in statistical sampling in the Indian statistical Institute, in: Journal of the Royal Statistical Society, Vol. 109, 1946, S. 325 - 370.
- Menges, G.: Stichproben aus endlichen Gesamtheiten, Theorie und Technik, Frankfurt 1959.
- Menges, G.: Grundriß der Statistik, Teil 1: Theorie, 2. Aufl. Opladen 1972.
- Menges, G.: Die statistische Adäquation, in: Jahrbücher für Nationalökonomie und Statistik, Bd. 197, 1982 a, S. 297 - 307.
- Menges G.: Die Statistik; Zwölf Stufen des statistischen Arbeitens, Wiesbaden 1982 b.
- Menges, G., Skala, H. J.: Grundriß der Statistik, Teil 2: Daten, Opladen 1973.
- Morgenstern, O.: Über die Genauigkeit wirtschaftlicher Beobachtungen, 2. Aufl., Wien, Würzburg 1965.
- Pokropp, F.: Stichproben: Theorie und Verfahren, Königstein 1980.
- Statistisches Bundesamt (Hrsg.): Stichproben in der amtlichen Statistik, Wiesbaden 1960.
- Stenger, H.: Stichprobentheorie, Würzburg 1971.
- Stenger, H.: Antwortvariabilität als Fehlerquelle in: Allgemeines Statistisches Archiv, Vol. 61, 1977, S. 391 - 403.

- Strecker, H.: Ein Modell zur Ermittlung von Angabefehlern in statistischen Erhebungen, in: *Metika* Vol. 17, 1971, S. 141 - 145.
- Strecker, H.: Zur Genauigkeit statistischer Daten, in: *Gegenwartsprobleme der Wirtschaft und der Wirtschaftswissenschaften*, hrsg. v. K. Born, Tübingen 1978, S. 85 - 115.
- Strecker, H.: Model for the Decomposition of Errors in Statistical Data into Components, and the Ascertainment of Respondent Errors by Means of Accuracy Checks, in: *Jahrbücher für Nationalökonomie und Statistik*, Vol. 195/5, 1980, S. 385 - 420.
- Strecker, H., Wiegert, R.: Erhebungsmodelle mit Wiederholungszählungen, in: *Messung der Antwortvariabilität auf Grund von Erhebungsmodellen mit Wiederholungszählungen*, hrsg. v. G. Tintner, H. Strecker, R. Férron, Göttingen 1983, S. 9 - 23.
- Strecker, H., Wiegert, R., Peeters, J., Kafka, K.: Erhebungssystem zur praktischen Bestimmung der Fehler und der Antwortvariabilität - bei Erhebungen in Belgien, Mai 1979, hrsg. v. G. Tintner, H. Strecker, R. Féron, Göttingen 1983, S. 24 - 112.
- Sukhatme, P. V., Sukhatme, B. V.: *Sampling Theory of Surveys with Applications*, 2. Aufl., London 1970.
- Warner, S. L.: Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias, in: *Journal of the American Statistical Association*, Vol. 60, 1965, S. 63 - 69.
- Wiegert, R.: Überlegungen zum Adäquations- und Fehlerbegriff in der Statistik und Ökonometrie, in: *Angewandte Statistik und Wirtschaftsforschung heute*, hrsg. v. W. Piesch u. W. Förster, Göttingen 1982, S. 254 - 266.

1 Einleitung

Viele Wissenschaftsdisziplinen benutzen statistische Methoden als Hilfsmittel, um Bereiche und Teilgebiete ihres Erkenntnisobjektes zu beschreiben oder Zusammenhänge empirisch aufzudecken. Außerdem sind theoretisch abgeleitete Interdependenzen an der Praxis zu überprüfen oder zu quantifizieren. Auch in den Wirtschafts- und Sozialwissenschaften spielen statistische Methoden eine bedeutende Rolle: Große Teile der wirtschaftswissenschaftlichen Forschung sind nur mit Anwendung statistischer Methoden denkbar. Wirtschaftspolitische Entscheidungen basieren nicht zuletzt auf Informationen, die mit statistischen Methoden ermittelt wurden.

Um die Ergebnisse der statistischen Erhebungen bei der Anwendung ihrer tatsächlichen Aussagefähigkeit entsprechend zu benutzen, sind Anhaltspunkte über deren Genauigkeit notwendig, denn jedes statistische Verfahren ist mit Fehlern behaftet¹⁾. Die Veröffentlichung von Ergebnissen statistischer Untersuchungen sollte daher Angaben über mögliche Fehler und Ungenauigkeiten enthalten²⁾. Dies schützt zum einen den Anwender der Ergebnisse vor Fehleinschätzungen und zum anderen wird dem Mißbrauch von statistischen Ergebnissen entgegengewirkt.

Die tatsächlich bestehenden Fehler einer statistischen Erhebung bleiben zwar unbekannt, aber Hinweise auf mögliche Ursachen und auf die mögliche Größe des statistischen Fehlers sind zur Beurteilung der Ergebnisse einer statistischen Untersuchung unabdingbar. In einigen Fachgebieten, insbesondere den Naturwissenschaften, ist die Angabe von Fehlergrenzen bei der Veröffentlichung von Meß- und Untersuchungsergebnissen üblich³⁾. Bereits vor

1) Vgl.: Menges u. Skala, 73, S. 272.

2) Vgl.: Statistisches Bundesamt, 60, S. 49; Anderson, 65, S. 89.

3) Vgl.: Kallmeyer, 56, S. 19; Morgenstern, 65, S. 6.

dreißig Jahren klagte O. Anderson darüber, daß sich dies für die Wirtschafts- und Sozialwissenschaften, insbesondere auch für die amtliche Statistik, noch nicht durchgesetzt hat¹⁾. Auch heute wird der Notwendigkeit, statistische Ergebnisse durch Angaben über ihre Genauigkeit zu kommentieren, "von der amtlichen Statistik nur unzureichend entsprochen"²⁾.

Als Begründung für diesen Mißstand wird meist angeführt, daß durch Fehlerangaben Mißverständnisse und Schwierigkeiten in der Öffentlichkeit entstünden³⁾. Dies mag damit zusammenhängen, daß die Ergebnisse statistischer Untersuchungen aus dem Bereich der Wirtschafts- und Sozialwissenschaften nicht ausschließlich für ein Fachpublikum bestimmt sind, das die Möglichkeiten der statistischen Methoden kennt und das mit Fehlern in den Ergebnissen umgehen kann. In der breiten Öffentlichkeit verhindert die Assoziation 'Fehler gleich falsch, ungenügend, unbrauchbar' den sachlichen Umgang mit statistischen Fehlern.

Von großer Bedeutung für die Beurteilung von Schätzergebnissen sind der Zufallsfehler und die Möglichkeiten, wie dieser aus der Erhebung zu ermitteln ist. Das Interesse der Statistiker richtete sich lange Zeit darauf, diesen Fehler durch Verbesserungen, Verfeinerungen der Verfahren zu reduzieren. Nach dem 2. Weltkrieg erstreckte sich die Aufmerksamkeit der Forscher auf dem Gebiet der Stichprobentheorie zunehmend auf Unzulänglichkeiten und Fehler, die bei der Durchführung der Erhebung fast zwangsläufig auftreten⁴⁾.

Im Bereich der Wirtschafts- und Sozialwissenschaften spielen Befragungen von Untersuchungseinheiten eine bedeutende Rolle, denn der Mensch mit seinem Wirtschaft-

1) Vgl.: Anderson, 54, S. 38 f.

2) Krug u. Nourney, 82, S. 183.

3) Vgl.: Kallmeyer, 56, S. 19; Morgenstern, 65, S. 9 f.

4) Vgl.: Strecker, 80, S. 385 f; Strecker u. Wiegert, 83, S. 9.

ten und Handeln steht oft im Mittelpunkt der Betrachtungen. In diesem Zusammenhang sind Fehler in den Antworten der Befragten besonders zu beachten. Wichtige Impulse zur theoretischen und empirischen Untersuchung der Qualität statistischer Daten gingen von Hansen und Hurwitz (U. S. Bureau of Census) und Mahalanobis (Indian Statistical Institute) aus¹⁾. H. Strecker hat mit seinen theoretischen und empirischen Arbeiten zur Messung von Antwortfehlern viel für die Verbreitung dieser Aspekte der Stichprobenerhebung im deutschsprachigen Raum getan.

Ziel dieser Arbeit ist es, die Fehler und Fehlerquellen bei der Durchführung einer Stichprobenerhebung genauer zu beschreiben und die Auswirkungen der Fehler auf die Schätzung, speziell die Mittelwertschätzung, zu untersuchen. Die Betrachtungen beschränken sich auf das Schichtenauswahlverfahren, wobei das Verfahren der einfachen Zufallsauswahl zur einfacheren Beschreibung der geschichteten Auswahl mit einbezogen ist. Dies ermöglicht gleichzeitig einen Vergleich beider Verfahren unter Berücksichtigung von verschiedenen Fehlern in der Erhebung.

Bevor die Auswirkungen der Fehler auf die Schätzung bei beiden Verfahren beleuchtet werden können, ist es sinnvoll, die möglichen Fehlerquellen einer Erhebung näher zu betrachten. Es wird sich herausstellen, daß für die Beschreibung der Auswirkungen von Fehlern auf die Schätzaussage entscheidend ist, ob diese mit den Antworten der befragten Einheiten verbunden sind, oder ob sie trotz korrekter Antworten bestehen. Im Zusammenhang mit den Antwortfehlern ist auf die Abgrenzung der Begriffe Zufallsfehler, systematischer Fehler und Verzerrung näher einzugehen.

In Kapitel 3 werden die Fehler in einer Erhebung bei

1) Vgl.: Stenger, 77, S. 392; Strecker, 80, S. 386;
Strecker u. Wiegert, 83, S. 9; Koop, 74,
S. 19 f.

korrekten Antworten näher betrachtet. Dazu ist eine kurze Beschreibung der einfachen Zufallsauswahl und des Schichtenauswahlverfahrens notwendig. Die Auswirkungen der aus den systematischen Fehlern in der Erhebung resultierenden Verzerrung auf die Konfidenzaussage bildet den Ausgangspunkt eines Vergleichs zwischen einer verzerrten und einer unverzerrten Schätzung. Es läßt sich ermitteln, welcher Stichprobenumfang ausgereicht hätte, ein Konfidenzintervall mit äquivalenter Aussage anzugeben, wenn die Verzerrung der Schätzung vermieden werden könnte. In Tabelle 3.4 sind die Ergebnisse eines solchen Vergleichs zusammengestellt.

Beim Schichtenauswahlverfahren wird der Schichteneffekt mit in die Betrachtungen der Auswirkungen einer Verzerrung auf die Schätzaussage einbezogen. Dadurch läßt sich ableiten, wie sich die Verbesserung der Schätzung durch die Schichtung bei gegebener Verzerrung auf die Konfidenzaussage auswirkt. Die Resultate der Berechnungen sind in Tabelle 3.6 angegeben.

Die Antwortfehler der Untersuchungseinheiten, die über einen zufälligen Fehlerterm zu beschreiben sind, werden in Kapitel 4 untersucht. Die Aufmerksamkeit gilt dabei besonders der Varianz des Mittelwertschätzers bei zufälligen Antworten und wie diese aus der Stichprobe zu schätzen ist. Durch die Fehler in den Antworten steigt der Zufallsfehler der Schätzung. Dies muß bei einer Gegenüberstellung zwischen einer verzerrten und einer unverzerrten Schätzung ebenso berücksichtigt werden, wie beim Vergleich zwischen einfacher Zufallsauswahl und geschichteter Auswahl.

Auf eine detaillierte Beschreibung der Auswirkungen einer durch die Antwortfehler entstandenen Verzerrung kann verzichtet werden, denn die Ergebnisse der Überlegungen aus Kapitel 3 lassen sich auf die Schätzung bei Antwortfehlern übertragen.

2 Fehler in statistischen Untersuchungen

2.1 Fehlerquellen in statistischen Untersuchungen

2.1.1 Überblick

Die Anwendung statistischer Verfahren in den Wissenschaftsdisziplinen sind jeweils auf ein bestimmtes Ziel, das Untersuchungsziel, ausgerichtet. Die Abgrenzung der Gesamtheit und die relevanten Untersuchungsmerkmale sind aus der Zielsetzung abzuleiten. Häufig sind Gegebenheiten und Zusammenhänge empirisch an Hand der Häufigkeitsverteilung(en) eines Merkmals (oder mehrerer Merkmale) in der Gesamtheit zu untersuchen. Oft begnügt sich der Forscher auch mit einem oder mehreren Parameter(n) dieser Häufigkeitsverteilung(en) wie Mittelwert oder Anteil der Merkmalsträger mit bestimmten Eigenschaften.

In dieser Arbeit gehen wir von einem Parameter¹⁾ Θ aus, der mit einer auf einer Zufallsauswahl beruhenden Stichprobe geschätzt werden soll. Bezogen auf diese Parameterschätzung - Fehlerbetrachtungen im Rahmen der Testtheorie: Fehler erster und zweiter Art, bleiben hier außer acht - kann der statistische Fehler als "die Abweichung des Ergebnisses einer Beobachtung oder eines Experiments von seinem 'wahren' Wert"²⁾ bezeichnet werden. Die Bemühungen des Statistikers müssen darauf gerichtet sein, diesen realisierten tatsächlichen Fehler³⁾ unter den gegebenen Umständen der Untersuchung,

-
- 1) Es sei unterstellt, daß dieser wahre Wert des Parameters existiert.
 - 2) Menges u. Skala, 73, S. 272; im Original hervorgehoben. Vgl. auch Krug u. Nourney, 82, S. 192. In der Aufzählung der statistischen Untersuchungen könnte auch die Befragung erwähnt werden.
 - 3) Manchmal auch als Gesamtfehler bezeichnet, vgl. Statistisches Bundesamt, 60, S. 20. Die Verwendung des Begriffs Gesamtfehler ist jedoch nicht einheitlich, vgl. Menges u. Skala, 73, S. 286.

(insbesondere der Kosten und eventuell zusätzlicher Aufgaben) möglichst gering zu halten¹⁾).

Wegen des Zufallscharakters der Stichprobe ist das Ergebnis der Untersuchung, der Schätzwert, die Realisation einer Zufallsvariablen $\hat{\theta}$. Der tatsächliche Fehler der Parameterschätzung ist ebenfalls Realisation einer Zufallsvariablen

$$T = \hat{\theta} - \theta, \quad (2.1)$$

die zur Beschreibung des statistischen Fehlers geeignet ist.

Der wahre Wert des zu schätzenden Parameters ist in der Regel unbekannt - andernfalls ist eine statistische Untersuchung zur Bestimmung von θ unnötig - und damit auch der realisierte tatsächliche Fehler. Lediglich das Ergebnis der Untersuchung, die Realisation von $\hat{\theta}$, ist bekannt. Um trotzdem Angaben über den Fehler machen zu können, soll im Rahmen jeder Erhebung eine Fehlerbetrachtung durchgeführt werden.

Ein umfassendes Bild über die möglichen Fehler in statistischen Untersuchungen ist durch eine detaillierte Betrachtung der Entstehungsursachen der Fehler zu erhalten. In der Literatur sind umfangreiche Arbeiten aus diesem Bereich zu finden, die zum Teil auch an konkreten Beispielen die verschiedenen Fehlerursachen aufzeigen²⁾. Das Hauptaugenmerk ist dabei auf Fehlerquellen gerichtet, in denen Mängel in der Erhebung ihren Ausdruck finden.

Eine umfangreiche Arbeit, die viel Beachtung gefunden hat, stammt von O. Morgenstern³⁾. Darin werden

1) Vgl.: Hansen u. a., I, 53, S. 56; Menges u. Skala, 7, S. 272.

2) Die im folgenden erwähnte Literatur stellt eine Auswahl aus der Fülle der Arbeiten dar; sie erhebt keinen Anspruch auf Vollständigkeit.

3) Vgl.: Morgenstern, 65; im engl. Original von 1950.

verschiedene Fehlermöglichkeiten aufgezeigt und einige Wirtschaftsstatistiken auf Fehler untersucht. Morgenstern weist auch auf die hohe 'Scheingenaugigkeit' der veröffentlichten Ergebnisse hin. Er nennt Schätzwerte, die mit vielen Stellen genau angegeben sind, obwohl bereits die dritte Stelle von links unsicher ist¹⁾.

G. Menges²⁾ gibt eine Liste mit 15 verschiedenen Fehlerquellen für Teilerhebungen an. Die Aufzählung zeigt die Vielfalt der Fehlerursachen auf, allerdings fehlen Hinweise auf Ähnlichkeiten und Strukturen in der Vielfalt der aus verschiedenen Quellen entstehenden Fehler.

Einige Autoren versuchen, die Fehlerquellen zu strukturieren. Bei Esenwein-Rothe³⁾ steht dabei die Entstehung von Fehlern im Vordergrund, um dann Erkenntnisse für die Vermeidung der daraus resultierenden Fehler zu gewinnen, die in Kontrollverfahren während der Erhebung ihre Anwendung finden (können).

Eine sehr klare Gliederung der Fehlerquellen geht auf eine Arbeit von Hurwitz und Pritzker zurück⁴⁾. Aufgegriffen und zum Teil modifiziert wurde die Einteilung u. a. von Strecker⁵⁾, Menges⁶⁾ und Krug und Nourney⁷⁾. In Übersicht 1 ist die Gliederung der Fehlerquellen nach Krug und Nourney in verkürzter und leicht abgewandelter Form wiedergegeben.

1) Vgl.: Morgenstern, 65, S. 61 ff; vgl. auch Anderson, 65, S. 80 f.

2) Vgl.: Menges, 59, S. 85 f; er bezieht sich darin auf Deming, 50, S. 26 ff. Vgl. auch Menges u. Skala, 73, S. 285 f.

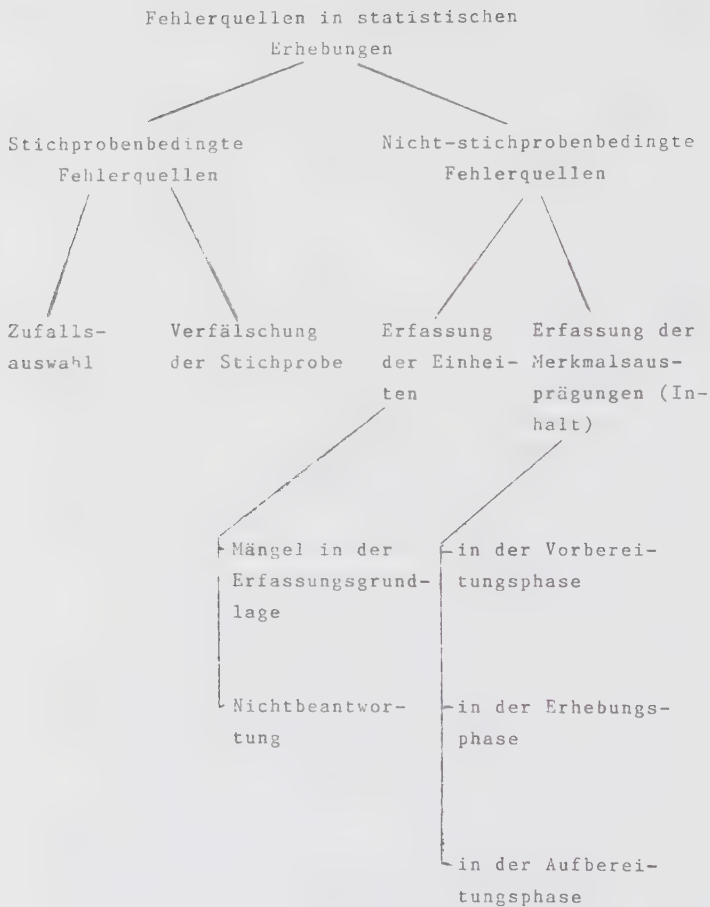
3) Vgl.: Esenwein-Rothe, 70, S. 167 ff.

4) Vgl.: Strecker, 80, S. 388f; Krug u. Nourney, 82, S. 159.

5) Vgl.: Strecker 80, S. 389.

6) Vgl.: Menges, 82 b, S. 450.

7) Vgl.: Krug u. Nourney, 82, S. 158 ff.



Übersicht 1: Mögliche Fehlerquellen in statistischen Erhebungen¹⁾

1) Vgl.: Krug u. Mourney, 32, S. 160.

Zum Teil sind die Gliederungen nicht auf die Fehlerquellen ausgerichtet, sondern auf die resultierenden Fehler. Da einige der möglichen Fehlerquellen nur indirekt auf die Qualität der Daten wirken, ergeben sich Unterschiede in der Aufgliederung.

Eine wichtige Unterscheidung der verschiedenen Erhebungen, auch im Hinblick auf die daraus resultierenden Fehlerquellen, ist die Einteilung in Totalerhebung und Teilerhebung. Entsprechend sind in Übersicht 1 die Fehlerquellen in zwei Gruppen gegliedert. In der ersten Gruppe sind die Entstehungsursachen der Fehler zusammengefaßt, die nur in Teilerhebungen auftreten können - die stichprobenbedingten Fehlerquellen. Die zweite Gruppe enthält Fehlerursachen, die nicht ausschließlich auf Teilerhebungen beschränkt sind, sondern auch in Totalerhebungen vorkommen - die nicht-stichprobenbedingten Fehlerquellen.

2.1.2 Stichprobenbedingte Fehlerquellen

Neben Mängeln und Unzulänglichkeiten in der Durchführung der Untersuchung bestehen bei Teilerhebungen Fehlermöglichkeiten, die mit der Auswahl der Untersuchungseinheiten aus der Gesamtheit zusammenhängen. Schon allein weil nur ein Teil der zur Gesamtheit gehörenden Elemente zur Analyse herangezogen wird, ist eine Abweichung der Ergebnisse vom wahren Wert möglich.

Wir beschränken uns hier auf Teilerhebungen, deren Elemente auf Grund einer Zufallsauswahl aus der Gesamtheit gezogen werden. Solche Stichproben lassen sich mit wahrscheinlichkeitstheoretischen Mitteln beschreiben. Der im Stichprobenergebnis bestehende Fehler resultiert aus dem Zufallsexperiment 'Stichprobe ziehen' und läßt sich deshalb als Zufallsvariable auffassen.

Außer diesem Zufallsfehler, der in der Literatur zur Stichprobentheorie eingehend beschrieben ist, können Stichprobenverfälschungen auftreten, einmal, weil das Stichprobenverfahren einen verzerrten Schätzer liefert, zum anderen, weil die Auswahl der Stichprobe mit Mängeln behaftet ist.

Die Gefahr der Verwendung verzerrter Schätzfunktionen, d. h. wenn der theoretische Erwartungswert des Schätzers nicht mit dem zu schätzenden Parameter der Gesamtheit übereinstimmt (sofern Fehler aus allen anderen hier aufgeführten Quellen ausgeschlossen sind), besteht bei den gebräuchlichen Stichprobenverfahren nicht. Wenn jedoch eine Anpassung des Verfahrens, auch durch Anwendung zusammengesetzter Methoden, an die jeweiligen Rahmenbedingungen der Untersuchung notwendig ist, können Fehler im Schätzer entstehen, die allerdings oft im Zusammenhang mit Fehlern in der Auswahlgrundlage - eine der drei möglichen Mängel bei der Auswahl der Stichprobe - zu sehen sind.

Die Fehler in der Auswahlgrundlage stehen in engem Zusammenhang mit den Erfassungsfehlern aus Gruppe zwei, denn Fehler bei der Abgrenzung der Grundgesamtheit (Erhebungsgesamtheit) und Mängel bei der Zusammenstellung der Einheiten in der Erfassungsgrundlage¹⁾ wirken sich auf die Liste²⁾ der Auswahlseinheiten, die als Grundlage für die Stichprobenauswahl dient, aus. Sind Auswahlseinheiten und Erhebungseinheiten identisch, schlagen sich Fehler in der Erfassungsgrundlage direkt in der Auswahlgrundlage nieder.

Die hier betrachteten Stichproben setzen eine reine Zufallsauswahl unter den in der Auswahlgesamtheit

1) Die Erfassungsgrundlage "besteht aus einem Verzeichnis von Einheiten in Form von Listen, Karteien, Datensätzen auf Magnetbändern oder auch in Form von Verwaltungseinheiten oder räumlich abgegrenzten Flächen. Die so verzeichneten Einheiten sind nicht notwendig mit den Erhebungseinheiten identisch, sondern können von weitgehend beliebiger, anderer Art sein". Krug u. Nourney, 82, S. 45.

2) Der Begriff ist nicht wörtlich zu nehmen, vgl.: 1).

zusammengefaßten Einheiten - wobei die Besonderheiten des jeweiligen Stichprobenverfahrens mit zu berücksichtigen sind - voraus. Die wegen ihrer einfacheren Handhabung häufig benutzten Ersatzverfahren¹⁾ für die reine Zufallsauswahl, die bei korrekter Anwendung der echten Zufallsauswahl gleichwertig sind, bergen die Gefahr einer Verzerrung des Auswahlverfahrens, z. B. durch "zyklische Besonderheiten"²⁾ in der geordneten Auswahlgrundlage. Zusätzlich führen Mängel in der Durchführung des an sich korrekten Auswahlverfahrens zu Fehlern in der Stichprobe.

2.1.3 Nicht-stichprobenbedingte Fehlerquellen

2.1.3.1 Fehlerquellen bei der Erfassung der Untersuchungseinheiten

Die wesentlich größere Zahl der Fehlermöglichkeiten in statistischen Untersuchungen kann man zur Gruppe der nicht-stichprobenbedingten Fehlerquellen zusammenfassen, die sowohl in Teil- als auch in Totalerhebungen auftreten können. Sie entstehen bei der Erfassung der Einheiten und bei der Feststellung der Merkmalsausprägungen.

Die Basis für die Durchführung einer Erhebung, unabhängig davon, ob es sich um eine Teil- oder eine Totalerhebung handelt, ist die Zusammenfassung der zur Gesamtheit gehörenden Untersuchungseinheiten in der Erfassungsgrundlage. Sie entspricht nicht den Erfordernissen, wenn Elemente der Gesamtheit nicht bekannt, doppelt erfaßt oder Einheiten, die nicht zur Erhebungsgesamtheit gehören, in ihr enthalten sind³⁾.

1) Vgl.: Statistisches Bundesamt, 60, S. 30 ff;
Krug u. Nourney, 82, S. 90 ff.

2) Krug u. Nourney, 82, S. 159.

3) Zu den Erfassungserlern vgl. auch: Esenwein-Rothe, 70, S. 172 ff.

Ein Grund für diese 'Coverage-errors' ist die zeitliche Diskrepanz zwischen Entstehung der Erfassungsgrundlage und Erhebungszeitpunkt. Außerdem existiert in der Erhebungspraxis nur in seltenen Fällen eine Zusammenfassung der für das konkrete Erhebungsziel relevanten Erhebungseinheiten, vielmehr muß eine, die Gesamtheit vollständig überdeckende, Erfassungsgrundlage aus verschiedenen vorhandenen Unterlagen zusammengestellt werden. Die Identifizierung der Einheiten als zur Gesamtheit gehörig erfolgt über Merkmale und Kriterien, die zur Abgrenzung der Erhebungsgesamtheit notwendig sind. Fehler in diesen Unterlagen sowie Überschneidungen und unterschiedliche Abgrenzungen der Begriffe schlagen sich dann in der Erhebungsgrundlage nieder. Nur eingehende Kontrollen, die oft mit viel Mühe und hohen Kosten verbunden sind, ermöglichen ein Reduzieren dieser Fehler. Auch die Wahl des Stichprobenverfahrens und die der Auswahlseinheiten beeinflußt die möglichen Erfassungsfehler.

Krug und Nourney zählen auch Fehler, die entstehen, wenn Einheiten, die befragt werden sollen, keine Antwort geben wollen oder können - Nichtbeantwortung - zu den Erfassungsfehlern. Als Gründe sind für diese Art des Fehlers denkbar: nicht auffinden, nicht antreffen des zu Befragenden, Verlust der Antwortbögen oder eine bewußte Antwortverweigerung.

2.1.3.2 Fehlerquellen bei der inhaltlichen Erfassung der Merkmalsausprägungen

Neben den Erfassungsfehlern beinhaltet auch die Feststellung der Merkmalsausprägungen wichtige Fehlerquellen im Rahmen einer statistischen Untersuchung. Diese entstehen in der Vorbereitungs-, der Erhebungs- und in der Aufbereitungsphase der Untersuchung. Krug und Nourney nennen eine ganze Reihe von möglichen

Fehlerursachen in der Vorbereitungsphase, angefangen von der Konzeption der Erhebung über die Termin- und Ablaufplanung, die Fragebogengestaltung bis hin zur finanziellen und personellen Ausstattung.

Die Art der Fragebogengestaltung, Formulierung und Reihenfolge der Fragen bestimmt, je nach Untersuchungsgegenstand, wesentlich die Qualität der Erhebung. Gerade in diesem Bereich besteht die Gefahr, daß bewußt oder unbewußt durch Formulierungen, Aufbau des Fragebogens Antworten beeinflußt, verfälscht werden¹⁾. Fehler dieser Art korrespondieren auch mit solchen, die in der Erhebungsphase der Untersuchung entstehen, insbesondere Antwortvariabilität und Interviewereinfluß. Zusätzlich können Mängel in der Termin- und Ablaufplanung, in der Voruntersuchung sowie bei der Personal- und Finanzausstattung negativen Einfluß auf die Ergebnisse der statistischen Untersuchung ausüben.

Bevor ein Fragebogen entworfen werden kann, muß das oft wenig operational formulierte Untersuchungsziel einer Erhebung in der Vorbereitungsphase konkretisiert werden²⁾. Im einzelnen muß die Erhebungsgesamtheit abgegrenzt werden, wobei vor allem die sachliche Abgrenzung schwierig ist und deshalb auch eine häufige Fehlerquelle darstellt (insbesondere auch im Zusammenhang mit der Erstellung der Erhebungsgrundlage). Ein weiterer mit vielen Fehlermöglichkeiten behafteter Schritt ist die Umsetzung des Erhebungsziels in erfragbare Informationen.

Fehler dieser Art sind von grundlegender Natur. Gelingt es nicht, die der Zielsetzung der Untersuchung zugrun-

1) Vgl.: Holm, 75 b, S. 63 f; Atteslander, 75, S. 100 ff; v. Kirschhofer-Bozenhardt u. Kaplitza, 75, S. 93 ff.

2) Vgl.: Holm, 75 a, S. 14 ff.

deliegenden theoretischen Sachverhalte in statistische Begriffe wie Einheiten, Merkmalsausprägungen usw. umzusetzen, mißt die Erhebung also nicht das, was sie messen soll, hat das Ergebnis kaum Aussagekraft für das angestrebte Erhebungsziel. Menges benutzt für diese Fehlerquelle den Begriff 'Adäquation'¹⁾. "Da die Adäquation ... nur immer graduell, niemals vollständig gelingen kann, bleibt stets eine Diskrepanz zwischen dem theoretisch eigentlich Gemeinten und dem statistisch Gemessenen"²⁾.

An dieser Stelle stößt man an die Grenze der Definition des tatsächlichen Fehlers T nach (2.1): Der wahre Wert des zu schätzenden Parameters Θ bezieht sich auf das Ziel der Untersuchung; Θ ist durch die Merkmalsausprägungen der zur Erhebungsgesamtheit gehörenden Elemente bestimmt. Mit der Annahme, daß Θ existiert, ist exakt gegeben, was mit der Untersuchung geschätzt werden soll. Damit ist das Problem der Adäquation umgangen.

Der Adäquationsfehler läßt sich quantitativ nicht erfassen, eben weil sich keine Parameter definieren lassen, die das Erhebungsziel vollständig beschreiben. Dies ist nur in Bezug auf eine 'beste Lösung' des Adäquationsproblems möglich, wobei noch ungeklärt ist, wie die verschiedenen Lösungen verglichen bzw. als äquivalent oder optimal erkannt werden können³⁾. Hinzu kommt, daß oft mehrere Merkmale zur umfassenden Beschreibung des Erhebungsziels notwendig sind, für die gemeinsam das Adäquationsproblem zu lösen ist.

1) Vgl.: Menges, 82 a, S. 39 ff.

2) Menges, 82 a, S. 302.

3) Vgl.: Wiegert, 83, S. 263 f.

Betrachtet man die anderen, auf die inhaltliche Erfassung der Merkmalsausprägung bezogenen, Fehlerquellen in der Vorbereitungsphase der Erhebung genauer, so stellt sich heraus, daß diese nur indirekt die statistische Untersuchung beeinflussen. Sie wirken sich erst in der Erhebungsphase, dann, wenn die Daten erhoben werden, aus.

Unzulänglichkeiten in der Fragebogengestaltung oder der Formulierung der Fragen können die Schätzung nur dann verschlechtern, wenn die Antwort des Befragten variabel ist, d. h., wenn Antwortvariabilität vorliegt¹⁾. Die Mängel in der finanziellen und personellen Ausstattung beeinflussen über den Interviewereinfluß oder die Nichtbeantwortung die Schätzaussage. Die Fehlerquellen in der Vorbereitungsphase einer Erhebung haben kein Pendant in einer entsprechenden Aufstellung der Fehler einer statistischen Untersuchung sofern man den Fehler quantitativ definiert. Aus diesem Grund sind sie auch nicht in der Aufstellung der Fehler nach Hurwitz und Pritzker enthalten²⁾, obwohl bei jeder statistischen Untersuchung diesen Fehlerquellen viel Aufmerksamkeit gebührt.

Fehler im Bereich der direkten Datengewinnung (Erhebungsphase) entstehen durch die Auskunftsperson, den Interviewer oder auch bei der Übertragung und Aufzeichnung der Daten. Letzteres ist durch Kontrollmaßnahmen und Sorgfalt weitgehend zu vermeiden. Falschantworten der Befragten treten sogar bei exakt bestimmten Informationen wie Familienstand, Alter oder Schulbildung auf. Empirische Untersuchungen zeigen, daß bei wiederholten Befragungen z.T. bei über

1) Die Fehlerquelle 'Antwortvariabilität' wird im Zusammenhang mit den Fehlern in der Erhebungsphase noch eingehender betrachtet.

2) Vgl.: Strecker, 80, S. 389.

20 % der Befragten die jeweils gegebenen Antworten nicht miteinander vereinbar waren¹⁾. Kallmeyer führt ein Beispiel aus der Agrarstatistik an, bei dem nur 2 % der Angaben über die Bodennutzung richtig waren; 46 % der Angaben waren zu hoch und 52 % zu niedrig²⁾.

Ein bekanntes Beispiel für solche Angabefehler ist der sogenannte Jubiläumseffekt, d. h., die quantitativen Angaben der Untersuchungseinheiten sind z. T. aus Unkenntnis der genauen Merkmalsausprägung auf "Jubiläumsszahlen" (10, 25, 50, 100 usw.) gerundet³⁾. Der Effekt wurde zum ersten mal bei Altersangaben von Personen in Ländern ohne genaue Einwohnermeldestatistik festgestellt. Es häuften sich die Angaben mit runden Zahlen. Strecker gibt eindrucksvolle Beispiele aus der Agrarstatistik⁴⁾.

Die Ursachen für die Falschantworten sind vielfältig und natürlich vom Untersuchungsgegenstand mit beeinflusst. Ein wesentlicher Aspekt ist die bewußte Falschantwort: Im Beispiel aus der Agrarstatistik, das Strecker erwähnt, beträgt der Anteil dieser Fehlerursache an den Falschantworten über 50 %⁵⁾, wobei noch die Gründe für diese bewußten Falschantworten - abhängig vom konkreten Erhebungsgegenstand - zu berücksichtigen sind.

Die Antworten der Befragten hängen von verschiedenen Einflüssen aus dem direkten Umfeld des Antwortenden und der Person des Antwortenden selbst ab. Man spricht in diesem Zusammenhang von Antwortvariabilität⁶⁾. Neben der Antwortvariabilität beeinflusst der Inter-

1) Vgl.: Stenger, 77, S. 391.

2) Vgl.: Kallmeyer, 56, S. 29.

3) Vgl.: Strecker, 78, S. 109; Anderson, 65, S. 80.

4) Vgl.: Strecker, 78, S. 109 f.

5) Vgl.: Strecker, 78, S. 108; Strecker u. a., 83, S. 38.

6) Vgl.: Stenger, 77, S. 391 f; Pokropp, 80, S. 201 ff.

viewereffekt, "d. h. eine durch die Subjektivität des Interviewers hervorgerufene Verfälschung der Beantwortung"¹⁾, die individuellen Angaben. Der Durchschnitt der Antworten, die bei einem Interviewer gegeben werden, ist verfälscht. Diese Verzerrung entsteht durch die Fehler in den individuellen Antworten, hervorgerufen vom Auftreten des Interviewers oder der Interaktion zwischen Interviewer und seinem Gegenüber.

Entscheidenden Einfluß auf den Interviewereffekt hat die Ausbildung der Befragenden, aber es bleiben auch bei bestem Ausbildungsstand Einflüsse bestehen²⁾.

Die Art und Größe des Interviewereinflusses hängt naturgemäß auch von der Art der Fragen ab. Werden Fakten erfragt, ist eine geringere Antwortverfälschung zu erwarten als bei Meinungsäußerungen und Beurteilungen durch den Befragten. Dies zeigen auch Ergebnisse von Untersuchungen über den Interviewereffekt³⁾.

Ein Interviewereinfluß setzt Antwortvariabilität voraus, denn nur wenn die Antwort des Befragten variabel ist, kann der Interviewer Einfluß auf die Antworten nehmen. Man kann deshalb den Interviewereinfluß als eine, allerdings besonders wichtige, Ursache für Falschantworten ansehen.

Streng genommen faßt der Ansatz der Antwortvariabilität eine Fülle verschiedener Einflußfaktoren unbekannter Stärke und Richtung zu einer Fehlervariablen zusammen. In der konkreten Erhebungssituation beeinflussen z. B. die persönlichen 'Launen' des Befragten, Informationen aus den Massenmedien, das Umfeld der Befragungssituation

1) Menges u. Skala, 73, S. 131.

2) Vgl.: Menges u. Skala, 73, S. 131.

3) Vgl.: Sukhatme u. Sukhatme, 70, S. 409 f. Sie beziehen sich dabei auf eine Untersuchung von Stock u. Hochstim aus dem Jahr 1951.

oder auch die Fragebogengestaltung die Antwort. Wegen der großen Bedeutung nimmt der Interviewereinfluß eine Sonderstellung unter den Ursachen der Antwortvariabilität ein. Er setzt die Möglichkeit einer variablen Antwort voraus, jedoch ist der Einfluß des Interviewers so groß, daß eine gesonderte Behandlung dieser Fehlerquelle angebracht sein kann.

Die Fehler in der Datengewinnung erhalten durch die Vielfalt der sich möglicherweise auch gegenseitig ausgleichenden Einflußfaktoren unter Umständen Zufallscharakter. Der so entstehende Zufallsfehler ist jedoch streng von dem oben erwähnten stichprobenbedingten Zufallsfehler zu trennen¹⁾.

Im Bereich der Datenaufbereitung sind neben Fehlern beim Codieren, Ablochen oder bei der Übertragung auf elektronische Datenträger auch fehlerhafte Berechnungen und Fehler beim Tabellieren der Ergebnisse zu beachten. Besondere Aufmerksamkeit ist auf die korrekte Interpretation der Untersuchungsergebnisse, insbesondere unter Beachtung möglicher Fehler in den Ergebnissen, zu richten²⁾.

1) Eine genauere begriffliche Trennung der verschiedenen Fehler erfolgt in Abschnitt 2.2.

2) Esenwein-Rothe spricht von Entscheidungsfehlern. Vgl.: Esenwein-Rothe, 70, S. 166.

2.2 Begriffliche Abgrenzung verschiedener Fehlerarten

2.2.1 Zufallsfehler, systematischer Fehler und Verzerrung

Aus der Aufstellung der Fehlerquellen geht hervor, daß die in statistischen Erhebungen auftretenden Fehler eine Vielzahl verschiedener Ursachen haben bzw. haben können. Dies erschwert eine Analyse der Auswirkungen und Versuche, die Größe und Richtung der Fehler anzugeben. Es ist nicht sinnvoll, die aus den verschiedenen Quellen stammenden Fehler einzeln zu betrachten, vielmehr soll versucht werden, nach ihren Auswirkungen auf die Schätzaussage Gruppen von Fehlern zu bilden, die gemeinsam behandelt werden können.

Üblicherweise wird in der Literatur zwischen Zufallsfehlern und systematischen Fehlern unterschieden. Leider ist die Abgrenzung dieser Fehlerarten nicht immer eindeutig bzw. einheitlich.

Einmal heißt es: "Die durch Stichproben bedingten Zufallsfehler ... sind nur ein - wenn auch bedeutsamer - Teil im gesamten System der möglichen Fehlerquellen in statistischen Ergebnissen. Den Zufallsfehlern steht nämlich eine größere Zahl andersartiger Fehler gegenüber, die unter dem Begriff systematische Fehler (im Original hervorgehoben, A. d. V.) zusammengefaßt werden"¹⁾. Nach dieser Definition sind alle aus den in Übersicht 1 genannten Quellen stammenden Fehler systematische, mit Ausnahme des stichprobenbedingten Zufallsfehlers. Die Bezeichnung systematische Fehler wird darin im Sinne von 'systematische Unzulänglichkeiten', als summarische, verbale Umschreibung für Mängel in einer statistischen Untersuchung benutzt.

1) Krug u. Nourney, 82, S. 157 f; vgl. auch Statistisches Bundesamt, 60, S. 20.

"Diese (systematischen A. d. V.) Fehler sind jedoch nicht so zu verstehen, daß sie je Art in jedem Einzelfall dieselbe Größe hätten, sondern sie wechseln in ihrer Größe und ggf. auch der Richtung, besitzen also eine Variabilität, die teilweise auch zufallsbestimmt sein kann¹⁾).

Eine andere begriffliche Abgrenzung zwischen Zufallsfehler und systematischem Fehler ist quantitativ ausgerichtet, auf die Abweichung einer Schätzung vom zu schätzenden Parameter²⁾).

Die Differenz des Schätzers $\hat{\theta}$ von seinem Erwartungswert $E(\hat{\theta})$

$$Z = \hat{\theta} - E(\hat{\theta}) \quad (2.2)$$

ist eine Zufallsvariable und verdient als solche die Bezeichnung Zufallsfehler.

Ist die Schätzung $\hat{\theta}$ nicht erwartungstreu, weil z. B. die Auswahlgrundlage verfälscht ist, oder weil Antwortfehler vorliegen, so besteht eine Verzerrung (bias)

$$b = E(\hat{\theta}) - \theta^3) \quad (2.3)$$

Zufallsfehler und Verzerrung ergeben zusammen den tatsächlichen Fehler, denn mit (2.2) und (2.3) folgt

$$\begin{aligned} 1 &= \hat{\theta} - \\ &= \hat{\theta} - E(\hat{\theta}) + E(\hat{\theta}) - \theta \\ &= Z + b. \end{aligned} \quad (2.4)$$

1) Krug u. Nourney, 82, S. 163.

2) Vgl.: Krug u. Nourney, 82, S. 192; Menges u. Skala, 78, S. 133 f.

3) Aus der Definition von b folgt $\theta = E(\hat{\theta}) - b$; den wahren Wert erhält man, wenn der Erwartungswert des Schätzers um b korrigiert wird (induktives Vorgehen), vgl.: Kallmeyer, 50, S. 19.

Für die Verzerrung einer Schätzung wird ebenfalls die Bezeichnung systematischer Fehler benutzt, das Pendant dazu ist der Zufallsfehler Z^1). "Im Sinne der zuvor getroffenen Einteilung in zufällige und systematische Fehler ist der zufällige Stichprobenfehler rein zufällig, die Stichprobenverfälschungen sind rein systematisch und die Nichtstichprobenfehler stellen Mischungen aus zufälligen und systematischen Fehlern dar"²).

Die beiden hier angeführten Abgrenzungen zwischen zufälligen und systematischen Fehlern weisen dann Unterschiede auf, wenn einige der Nichtstichprobenfehler Zufallscharakter annehmen. Dies ist in der Regel dann der Fall, wenn die Antworten der Befragten keine exakt definierten Werte sind - wenn Antwortvariabilität besteht. Einfluß übt auch die Fragebogengestaltung und der Interviewer aus, ex definitione aber nur dann, wenn die Antwort variabel sein kann.

Die Nichtbeantwortung kann ebenfalls zu einem Zufallsprozeß führen, z. B. dann, wenn die 'Entscheidung', ob der Befragte antwortet oder nicht, zufallsbestimmt ist.

Sind solche 'nichtstichprobenbedingten' Zufallskomponenten in einer Befragung auszuschließen, so ist der Zufallsfehler Z gleich dem stichprobenbedingten Zufallsfehler und entsprechend führen die systematischen Fehler zur Verzerrung b. Wenn man nicht begrifflich zwischen einer verbalen Bezeichnung der aus den verschiedenen Quellen stammenden Fehler und

1) Vgl.: Menges u. Skala, 73, S. 283; Krug u. Nourney, 82, S. 192.

2) Menges, 82 b, S. 451; Menges benutzt die Fehler-einteilung nach Strecker (vgl. Strecker, 80, S. 389), die in englischer Sprache verfaßt ist; daher entstehen Abweichungen in der Bezeichnung (vgl.: Menges, 82 b, S. 450). Die Zuordnung ist jedoch eindeutig.

dem quantitativen Fehler einer Schätzung trennen will, kann man dann auch die Verzerrung mit dem systematischen Fehler gleichsetzen.

In dieser Arbeit wird der Begriff 'systematische Fehler' als verbale Bezeichnung der Fehler in einer Erhebung, die auf Mängel und Unzulänglichkeiten in der Durchführung zurückzuführen sind, benutzt, d. h. systematische Fehler sind alle aus den in Übersicht 1 genannten Quellen stammenden Fehler, mit Ausnahme des stichprobenbedingten Zufallsfehlers.

Die Bezeichnung 'Verzerrung' bezieht sich ausschließlich auf eine konstante Abweichung des Erwartungswertes eines Schätzers vom zu schätzenden Parameter entsprechend Definition (2.3). Die systematischen Fehler führen zu der Verzerrung des Schätzers. Das Pendant zur Verzerrung ist der Zufallsfehler Z.

Wenn einzelne systematische Fehler in der Erhebung Zufallscharakter annehmen, erfaßt Z neben dem Zufallsfehler durch die Stichprobenauswahl auch diese Einflüsse. Man könnte daher zwischen dem stichprobenbedingten und dem nicht-stichprobenbedingten Zufallsfehler einer Schätzung unterscheiden¹⁾.

Eine exakte Trennung der beiden Arten von Zufallsfehlern läßt sich am Beispiel der Antwortfehler²⁾ vornehmen. Dazu ist eine formale Beschreibung der Stichprobenauswahl unter den Untersuchungseinheiten und des Antwortverhaltens der befragten Personen notwendig.

1) Auf die Möglichkeit einer solchen terminologischen Unterscheidung weisen auch Menges u. Skala hin. Vgl.: Menges u. Skala, 73, S. 278 f.

2) Andere systematische Fehler in der Erhebung, die u. U. Zufallscharakter annehmen können, bleiben im folgenden unberücksichtigt.

2.2.2 Die Stichprobenauswahl unter Berücksichtigung des Antwortverhaltens der Untersuchungseinheiten

2.2.2.1 Die Zufallsstichprobe bei korrekten individuellen Antworten

Gehen wir von einer endlichen Gesamtheit $G^{1)}$ mit N Untersuchungseinheiten u_j , $j = 1, 2, \dots, N$, aus:

$$G = \{u_j / j = 1, 2, \dots, N\}. \quad (2.5)$$

Jedem Element aus G ordnet das Untersuchungsmerkmal X eine Merkmalsausprägung

$$x_j = X(u_j), \quad x_j \in \mathbb{R}, \quad j = 1, 2, \dots, N, \quad (2.6)$$

zu. Wegen der Annahme, daß mit x_j die wahre Merkmalsausprägung von u_j existiert, ist das Problem der Adäquation - übertragen auf die Ebene der Merkmalsträger - umgangen. In der Regel wird der zu schätzende Parameter über die Merkmalsausprägungen x_j definiert²⁾: z. B. Mittelwert der quantitativen Merkmalsausprägungen der u_j aus G , also

$$\mu = \frac{1}{N} \sum_{j=1}^N x_j. \quad (2.7)$$

oder auch die Merkmalssumme $N \cdot \mu$. Bei qualitativen Da-

1) In der Praxis der Statistik im Bereich der Wirtschafts- und Sozialwissenschaften ist G meist endlich. Die Unterscheidung in endliche und unendliche Gesamtheiten ist für die theoretische Berechnung der Eigenschaften einer Schätzfunktion bedeutsam, auch wenn für genügend große endliche Gesamtheiten der Unterschied im praktischen Schätzergebnis vernachlässigbar bleibt.

2) Vgl.: Hansen u. a., 51, S. 149.

ten interessiert oft, wie groß die Anzahl bzw. der Anteil der Elemente mit einer bestimmten Eigenschaft in der Gesamtheit ist. Durch die Definition von $x_j = 0$, falls die Eigenschaft nicht erfüllt ist, und $x_j = 1$, falls die Eigenschaft erfüllt ist, bezeichnet μ den Anteil und $N \cdot \mu$ die Anzahl der Elemente mit $x_j = 1$ in G .

Im folgenden beschränken wir uns auf quantitative Merkmale, so daß wir μ als Mittelwert der Merkmalsausprägungen in G auffassen. Dieser Mittelwert ist durch die Stichprobe zu schätzen.

Die x_j und demzufolge auch μ als zu schätzender Parameter müssen dem Ziel der Untersuchung angemessen sein. Aus erhebungstechnischen Gründen sollte, wenn möglich, eine operationale Definition gewählt werden¹⁾, die unabhängig von den Erhebungsumständen ist.

In vielen Fällen ist es keine einschränkende Bedingung, daß die x_j existieren. Beispielsweise lassen sich exakte Werte angeben für das Einkommen oder die Konsumausgaben oder das Alter von Personen; ebenso ist das Investitionsvolumen oder der Umsatz eines Unternehmens im Jahr eine exakt definierte Größe.

Beschränken wir uns auf Stichprobenverfahren, bei denen Auswahlinheit und Untersuchungseinheit identisch sind, so kann man die Stichprobe vom Umfang n als Zufallsvektor

$$\underline{U} = (U_i / i = 1, 2, \dots, n) \quad (2.8)$$

auffassen²⁾. Die Menge der Realisationen der Zufalls-

1) Vgl.: Hansen u. a., 51, S. 150 f.

2) Vgl.: z. B. Pokropp, 80, S. 6. Es gibt noch eine andere Möglichkeit eine Stichprobe mit den Mitteln der Wahrscheinlichkeitstheorie zu beschreiben; vgl.: Cochran, 72, S. 225 ff. Die Stichprobe ist in diesem Ansatz eine Teilmenge aus G , vgl.: Pokropp, 80, S. 7.

variablen U_i ist jeweils die Gesamtheit G oder Teilmengen daraus, je nach dem konkret verwendeten Stichprobenverfahren. Die Verteilung der Zufallsvariablen ist mit dem Stichprobenverfahren ebenfalls festgelegt.

Können wir annehmen, daß die ausgewählten Einheiten mit ihren wahren Merkmalsausprägungen antworten, so beschreibt (2.6) das Antwortverhalten der Untersuchungseinheiten. Jeder Realisation von U_i ist eine Antwort zugeordnet.

Dies ermöglicht den Übergang von der Zufallsvariablen U_i , die die Auswahl aus den Untersuchungseinheiten beschreibt, zu einer Zufallsvariablen X_i , in der die Antworten der Einheiten berücksichtigt sind. Es gilt:

$$X_i = x_j \text{ falls } U_i = u_j; \quad (2.9)$$

die U_i bestimmen die Verteilung der X_i .

Die Stichprobe aus den Merkmalsausprägungen ist dann der Zufallsvektor

$$\underline{X} = (X_i / i = 1, 2, \dots, n). \quad (2.10)$$

Der Schätzer $\hat{\theta}_X^{(1)}$ ist eine Funktion von $\underline{X}$.

Erst durch die Befragung sind über n Antworten der ausgewählten Einheiten die reellen Merkmalsausprägungen in die Betrachtungen einzubeziehen. Zur stochastischen Beschreibung ist ein Antwortverhalten der Individuen zu unterstellen.

-
- 1) Die Annahme, daß die Untersuchungseinheiten mit ihren wahren Merkmalsausprägungen antworten, ist durch den Zusatz X hervorgehoben.

2.2.2.2 Die Antworten der Untersuchungseinheiten als Zufallsvariable

Im vorigen Abschnitt wurde mit dem Antwortverhalten (2.6) unterstellt, daß die Untersuchungseinheiten mit ihren wahren Merkmalsausprägungen x_j antworten. Die Ausführungen über die Fehlerquellen in Stichprobenerhebungen legen jedoch die Vermutung nahe, daß die bei der Untersuchung erhaltenen Daten nicht immer mit den wahren Merkmalsausprägungen der ausgewählten Einheiten übereinstimmen. Insbesondere Antwortvariabilität und Interviewereinfluß führen zu Fehlern in den Daten.

Im Zusammenhang mit den systematischen Fehlern ist zu beachten, daß sich die Untersuchungseinheiten nicht immer über ihre wahren Merkmalsausprägungen im klaren sind. In einer solchen Situation erhalten die Fehlerquellen, die zu Antwortfehlern führen, noch größere Bedeutung. Gleiches gilt auch dann, wenn sich die wahren Merkmalsausprägungen x_j nicht eindeutig aus sachlichen Definitionen ableiten lassen¹⁾.

Für die weiteren Betrachtungen setzen wir voraus, daß die x_j der Elemente aus der Gesamtheit G existieren. Das heißt nicht, daß die befragten Einheiten auch mit ihrem jeweiligen x_j antworten, denn die systematischen Fehler führen u. U. zu Falschantworten. Allerdings gehen wir davon aus, daß alle ausgewählten Untersuchungseinheiten eine Antwort geben. Das Problem der Nichtbeantwortung bleibt in den Betrachtungen unberücksichtigt.

Gehen wir von einer Antwort des Individuums u_j aus. Eine mehrfache Befragung des gleichen u_j führt

¹⁾ Vgl.: Pokropp, 80. S. 201 f.; Hansen u. v., 51. S. 149 ff.

wegen der o. g. Gründe nicht unbedingt zu gleichen Antworten. Wenn eine exakte Merkmalsausprägung der Einheit u_j mit x_j gegeben ist, läßt sich die Antwort y_{jk} bei der k -ten Befragung, $k = 1, 2, \dots, K$, aufspalten in¹⁾

$$\begin{aligned} y_{jk} &= x_j + e_{jk}; & (2.11) \\ e_{jk} &\in \mathbb{R}, k = 1, 2, \dots, K, K \geq 1, \\ x_j &\in \mathbb{R}, \text{konstant.} \end{aligned}$$

Die x_j sind durch einen Antwortfehler überlagert, der bei Befragung k die Ausprägung e_{jk} hat. In e_{jk} kommen die verschiedenen Fehlerquellen zum Tragen. Empirisch ist es unter Umständen sehr schwer, x_j und e_{jk} zu trennen, denn nur y_{jk} ist nach der Befragung bekannt.

Für die theoretische Behandlung der individuellen Antwortfehler ist eine solche deduktive Aufspaltung der Antwort besser geeignet als der induktive, auf die Korrektur von Ergebnissen ausgerichtete, Ansatz²⁾. Der individuelle Antwortfehler ist nicht konstant, vielmehr hängt er von vielen Faktoren ab; e_{jk} kann man daher als 'Realisation einer Zufallsvariablen ε_j^k ', 'Antwortfehler der Einheit u_j ', auffassen. Die Antwort des Individuums ist nach diesem Ansatz, unabhängig von der Stichprobenauswahl, selbst eine Zufallsvariable.

In der Fehlervariablen ε_j^k manifestieren sich die verschiedenen Einflüsse auf die Antwort der Untersuchungseinheit u_j , die in der Person des Befragten liegen,

1) Neben der in (2.11) zum Ausdruck kommenden additiven Überlagerung des wahren Wertes der Merkmalsausprägung durch den Fehlerterm sind auch andere Verknüpfungsarten, z. B. Multiplikation, denkbar. Für den Bereich der Wirtschafts- und Sozialwissenschaften sind solche Ansätze jedoch unüblich und der konkreten Erhebungssituation wenig angepaßt.

2) Vgl.: Kallmeyer, 56, S. 19; Strecker, 71, S. 142.

aber auch die systematischen Fehler in den Daten sind darin enthalten. Soll der Interviewereffekt in den Antworten unterschieden werden, so ist die Abhängigkeit der ϵ_j^i vom Interviewer besonders zu berücksichtigen¹⁾.

Die weiteren Betrachtungen in dieser Arbeit sind mehr auf die Antwortfehler als grundlegende Fehlerquelle ausgerichtet.

Strecker und andere haben in einer Untersuchung aus dem Bereich der Agrarstatistik (Schweinebestand) die individuellen Antwortfehler untersucht. Die Antworten wurden mit denen einer von besonders geschulten Interviewern durchgeführten Kontrollerhebung verglichen. Die Angabefehler schwankten zwischen -20 und +12 mit einem Mittelwert von -1,2 und einer Standardabweichung von 7,34²⁾.

Nach dem bisher gesagten läßt sich die Antwort der Individuen als Zufallsvariable Y_j^i auffassen, definiert mit

$$Y_j^i = x_j + \epsilon_j^i, \quad j = 1, 2, \dots, N. \quad (2.12)$$

Für die weitere Betrachtung müssen wir unterstellen, daß die Erwartungswerte und die Varianzen der Y_j^i bzw. ϵ_j^i existieren, d. h., "daß es für jedes $u_j \in G$ ein wohldefiniertes 'Durchschnittsverhalten' bei den Y -Antworten gibt"³⁾, d. h.

1) Vgl.: z. B. Sukhatme u. Sukhatme, 72, S. 390 ff.

2) Vgl.: Strecker 78, S. 107 f; zu beachten ist, daß die Angaben bei Strecker auf einem induktivem Fehleransatz beruhen, so daß für die hier gewählte Definition des Angabefehlers die Vorzeichen vertauscht sind. Der relative Angabefehler beträgt etwa 8%. Vgl. auch die Übersicht der Mittelwerte der Angabefehler der landwirtschaftlichen Betriebsfläche je Betrieb bzw. Farm in Strecker, 78, S. 87 f.

3) Pokropp, 80, S. 205.

$$E(\varepsilon_j) =: b_j, \quad j = 1, 2, \dots, N, \quad (2.13)$$

existiert.

Diesen erwarteten Antwortfehler der Einheit u_j kann man als individuelle Antwortverzerrung bezeichnen. Mit (2.13) ist der Erwartungswert der Antworten

$$\begin{aligned} E(Y_j') &= x_j + E(\varepsilon_j') \\ &= x_j + b_j \end{aligned} \quad (2.14)$$

additiv aus der wahren Merkmalsausprägung und den individuellen Verzerrungen der u_j zusammengesetzt. Da die Varianz einer Zufallsvariablen translationsinvariant ist, gilt:

$$\begin{aligned} \text{Var}(Y_j') &= \text{Var}(x_j + \varepsilon_j') \\ &= \text{Var}(\varepsilon_j') = \sigma_j^2. \end{aligned} \quad (2.15)$$

Bisher wurde immer unterstellt, daß ein wohldefinierter wahrer Wert der Merkmalsausprägung der Einheit u_j existiert. In vielen Fällen ist dies auch keine einschränkende Bedingung, z. B. bei soziodemographischen Merkmalen.

Es gibt jedoch auch Untersuchungsgegenstände, insbesondere aus dem Bereich der Sozialwissenschaften, bei denen den Einheiten u_j nicht eindeutig eine Merkmalsausprägung zugeordnet werden kann, der wahre Wert x_j ist nur schwer zu definieren (z. B. Meinungen zu bestimmten Problemen).

Können wir trotz dieser Schwierigkeiten ein Antwortverhalten der Individuen unterstellen, so besteht die Möglichkeit, den Erwartungswert der Zufallsvariablen 'Antwort der Einheit u_j ' als wahre Merkmalsausprägung zu definieren: Man setzt also

$$x_j := E(Y_j'), \quad j = 1, 2, \dots, N. \quad (2.16)$$

Unter dieser Voraussetzung gilt dann

$$E(\varepsilon_j') = 0,$$

d. h. die Antwort der Einheiten ist nicht verzerrt, aber variabel¹⁾.

Ein anderer Extremfall der Zufallsvariablen Y_j' ist noch zu erwähnen: $\sigma_j^2 = 0$, d. h., ε_j' ist einpunktverteilt. Für das Antwortverhalten bedeutet das eine verzerrte, aber keine variable Antwort.

Im Anschluß an das Ziehen der Stichprobe $\underline{U}$ aus der Menge der Untersuchungseinheiten erfolgt die Feststellung der Merkmalsausprägungen. Über die Zufallsvariablen Y_j' , $j = 1, 2, \dots, N$, die das Antwortverhalten der Merkmalsträger beschreiben, ist jeder Realisation u_j eine Antwortvariable zugeordnet. Analog zu (2.9) kann man daher Zufallsvariablen Y_i , $i = 1, 2, \dots, n$, definieren mit

$$Y_i = Y_j' \text{ falls } u_i = u_j \quad (2.17)$$

gilt. In den Y_i ist sowohl die Stichprobenauswahl als auch der Antwortfehler der ausgewählten Einheit berücksichtigt. Die Stichprobe aus den Antwortvariablen führt zum Zufallsvektor

$$\underline{Y} = (Y_i / i = 1, 2, \dots, n). \quad (2.18)$$

Der Schätzer $\hat{\theta}_Y^{(2)}$ ist eine Funktion dieses Zufallsvektors. Die Verteilung von $\hat{\theta}_Y$ ist natürlich von den Fehlern in den Antworten beeinflusst. Sowohl $E(\hat{\theta}_Y)$ als auch $\text{Var}(\hat{\theta}_Y)$ enthalten Komponenten der Antwortfehler.

1) Vgl.: Pokropp, 80, S. 204.

2) Die Annahme, daß die Antworten der Untersuchungseinheiten durch einen zufälligen Fehlerterm überlagert sind, kommt in dem Zusatz Y zu Ausdruck.

2.2.3 Der Einfluß der Antwortfehler auf den Zufallsfehler und die Verzerrung der Schätzung

Die Schätzer $\hat{\theta}_X$ bzw. $\hat{\theta}_Y$ sind als Funktionen von Zufallsvariablen ebenfalls Zufallsvariablen und haben eine Verteilung mit $E(\hat{\theta}_X)$ bzw. $E(\hat{\theta}_Y)$ und der Varianz $\text{Var}(\hat{\theta}_X)$ bzw. $\text{Var}(\hat{\theta}_Y)$. Wir unterstellen, daß diese Parameter existieren, was für die Situation der Stichprobentheorie keine einschränkende Bedingung ist.

In Analogie zu (2.2), (2.3) und (2.4) können wir für die Schätzer jeweils den Zufallsfehler, die Verzerrung und den tatsächlichen Fehler definieren:

$$Z_X = \hat{\theta}_X - E(\hat{\theta}_X); \quad (2.19)$$

$$b_X = E(\hat{\theta}_X) - \theta; \quad (2.20)$$

$$T_X = Z_X + b_X; \quad (2.21)$$

$$Z_Y = \hat{\theta}_Y - E(\hat{\theta}_Y); \quad (2.22)$$

$$b_Y = E(\hat{\theta}_Y) - \theta; \quad (2.23)$$

$$T_Y = Z_Y + b_Y. \quad (2.24)$$

Der Einfluß der Antwortfehler auf die Schätzung θ_Y läßt sich durch die Differenz¹⁾

$$\hat{\theta}_Y - \hat{\theta}_X = \hat{\theta}_E \quad (2.25)$$

beschreiben; daher ist die Bezeichnung $\hat{\theta}_E$ angebracht. Analog gilt für die Erwartungswerte

$$E(\hat{\theta}_Y) - E(\hat{\theta}_X) = E(\hat{\theta}_E). \quad (2.26)$$

1) Es ist unterstellt, daß bei beiden Schätzern die gleichen systematischen Fehler auftreten, mit Ausnahme der Antwortfehler. Der funktionale Zusammenhang zwischen Zufallsvektor und Schätzer ist ebenfalls gleich.

Der Zufallsfehler Z_Y läßt sich mit (2.25) und (2.26) zerlegen in

$$Z_Y = Z_X + Z_\varepsilon, \quad (2.27)$$

$$Z_\varepsilon = \hat{\theta}_\varepsilon - E(\hat{\theta}_\varepsilon).$$

In Z_X sind die Auswirkungen der Zufallsauswahl unter den wahren Merkmalsausprägungen der Untersuchungseinheiten aus G zusammengefaßt - Z_X beschreibt den stichprobenbedingten Zufallsfehler; Z_ε erfaßt den Einfluß der Antwortfehler auf die Schätzung - Z_ε ist der nicht-stichprobenbedingte Zufallsfehler.

Bei dieser Einteilung ist zu beachten, daß Z_ε nicht unbeeinflußt von der Stichprobenauswahl bleibt, denn $\hat{\theta}_Y$ enthält auch den Einfluß der Zufallsauswahl auf die individuellen Verzerrungen b_j .

Man könnte die Abgrenzung der beiden Zufallsfehler daher auch anders wählen und nur den Einfluß der Antwortvariabilität ε_j^1 auf den Schätzer $\hat{\theta}_Y$ als nicht-stichprobenbedingt bezeichnen. Der stichprobenbedingte Zufallsfehler wäre dann weiter gefaßt: er enthielte sowohl Z_X als auch den Teil von Z_ε , der durch die b_j bei der Zufallsauswahl entsteht¹⁾.

Da die b_j aber aus einer nicht-stichprobenbedingten Fehlerquelle stammen, dem Antwortfehler, sollen sie hier, trotz des Einflusses der Zufallsauswahl auf Z_ε , nicht zu den stichprobenbedingten Zufallsfehlern gerechnet werden; es gilt die Abgrenzung nach (2.27).

1) Einige Autoren benutzen diese Abgrenzung, wobei sie jedoch nicht explizit eine Aufspaltung des Zufallsfehlers vornehmen. Vgl.: Stenger, 77, S. 393 f.; Cochran, 72, S. 439 ff.

Analog zum Zufallsfehler setzt sich die Verzerrung b_Y der Schätzung $\hat{\Theta}_Y$ zusammen aus einem Teil, der durch die Antwortfehler entsteht, und einem auf Grund anderer systematischer Fehler. Durch Einsetzen von (2.26) und (2.20) in (2.23) folgt

$$\begin{aligned} b_Y &= E(\hat{\Theta}_X) + E(\hat{\Theta}_E) - \Theta \\ &= b_X + E(\hat{\Theta}_E). \end{aligned} \quad (2.28)$$

Der tatsächliche Fehler des Schätzers läßt sich mit (2.27) und (2.28) in vier Komponenten zerlegen, wobei wir für $E(\hat{\Theta}_E) = b_E$ schreiben:

$$T_Y = Z_X + Z_E + b_X + b_E, \quad (2.29)$$

Die bisherigen Betrachtungen zeigen, daß das Antwortverhalten der Individuen große Bedeutung für die Beschreibung der Fehler in Stichprobenuntersuchungen hat. Unter diesem Blickwinkel ist eine Zweiteilung der systematischen Fehler angebracht:

- (1) systematische Fehler in der Schätzung, bei korrekten individuellen Antworten (Antwortverhalten (2.6));
- (2) Antwortfehler der Untersuchungseinheiten (Antwortverhalten (2.12)).

Mit dem Antwortverhalten (2.6) lassen sich die Auswirkungen der stichprobenbedingten Zufallsfehler Z_X und die Verzerrung b_X der Schätzung als Ganzes auf die Aussage der Erhebung beschreiben. Dies soll in Kapitel 3 geschehen.

Anschließend, in Kapitel 4, werden systematische Fehler in den Antworten zugelassen, so daß der Zufallsfehler Z_E und die Verzerrung b_E Einfluß auf die Schätzung haben. Zur besseren Darstellung der

Antwortfehler ist es sinnvoll anzunehmen, daß keine systematischen Fehler auftreten, die zu einer Verzerrung b_X führen.

Die quantitativ nur unzureichend erfaßbaren Adäquations- und Interpretationsfehler sind unberücksichtigt - was nicht heißt, daß sie für statistische Erhebungen von untergeordneter Bedeutung sind. Dieses Vorgehen ist gerechtfertigt, weil wir in (2.6) bzw. (2.12) die adäquate Formulierung des Erhebungsziels der Untersuchung durch die wahren Merkmalsausprägungen der u_j voraussetzen müssen. Bei endlicher Gesamtheit ist dann auch der zu schätzende Parameter gegeben.

Die Interpretationsfehler entziehen sich weitgehend einer statistischen Beschreibung. Natürlich können zufällige oder systematische Fehler in der Erhebung solche Fehlinterpretationen auslösen; umso wichtiger ist daher die Betrachtung der Fehler einer statistischen Erhebung.

2.2.4 Tatsächlicher Fehler und Mean Square Error

Ausschlaggebend für die Aussagefähigkeit des Ergebnisses einer statistischen Untersuchung ist der tatsächliche Fehler, der sich additiv aus dem Zufallsfehler Z_Y und der Verzerrung b_Y zusammensetzt¹⁾ (vgl. (2.1) b. w. (2.24)):

$$\begin{aligned} \Gamma_Y &= \hat{\Theta}_Y - \Theta \\ &= Z_Y + b_Y. \end{aligned}$$

1) Es ist sinnvoll, den Schätzer $\hat{\Theta}_Y$ zu betrachten, denn er umfaßt neben den systematischen Fehlern, die bei korrekten Antworten der Untersuchungseinheiten auftreten, auch die Antwortfehler.

Die Definition des Zufallsfehlers bedingt, daß der Erwartungswert von Z_Y gleich 0 ist:

$$E(Z_Y) = E(\hat{\Theta}_Y - E(\hat{\Theta}_Y)) = 0. \quad (2.30)$$

Es ist zu beachten, daß die Erwartungswertbildung sowohl über die Stichprobenauswahl als auch über das Antwortverhalten der Individuen - die Zufallsvariable Y_j - erfolgen muß.

Für den tatsächlichen Fehler, der ja wegen Z_Y ebenfalls eine Zufallsvariable ist, folgt daraus

$$E(T_Y) = E(Z_Y) + E(b_Y) = b_Y, \quad (2.31)$$

denn die Verzerrung b_Y ist bei gegebenem Erhebungssystem¹⁾ konstant.

Da sich bei der Berechnung des Erwartungswertes von T_Y die Zufallsfehler aufheben, eignet sich $E(T_Y)$ nicht als Maß für den tatsächlichen Fehler. Um zu verhindern, daß sich positive und negative Abweichungen bei der Erwartungswertbildung aufheben, bietet sich $E(T_Y^2)$ an. Durch die Quadrierung des Fehlers erfolgt eine stärkere Gewichtung der absolut größeren und eine Verkleinerung der ganz kleinen absoluten Abweichungen.

Mit der Definition von T_Y nach (2.24) gilt²⁾

$$\begin{aligned} E(T_Y^2) &= E((\hat{\Theta}_Y - \Theta)^2) \\ &= E((Z_Y + b_Y)^2) \\ &= E(Z_Y^2) + 2E(Z_Y b_Y) + E(b_Y^2) \end{aligned}$$

und, da b_Y konstant ist, folgt mit (2.30) weiter

1) Das Erhebungs- oder Arbeitssystem beschreibt die Datenbeschaffung bei der Erhebung. Zum Begriff des Erhebungssystems vgl. Strecker, 78, S. 86.

2) Vgl.: z. B. Menges u. Skala, 73, S. 286; Stenger, 77, S. 393 f.

$$\begin{aligned}
 E(T_Y^2) &= E(Z_Y^2) + 2 b_Y E(Z_Y) + b_Y^2 \\
 &= E(Z_Y^2) + b_Y^2 \\
 &= \text{Var}(Z_Y) + b_Y^2.
 \end{aligned}
 \tag{2.32}$$

Der mittlere quadratische Fehler der Schätzung¹⁾ setzt sich additiv aus der Varianz $\text{Var}(Z_Y)$ des Zufallsfehlers, die gleich der des Schätzers $\hat{\Theta}_Y$ ist, und dem Quadrat der Verzerrung zusammen. Oft wird für $E(T_Y^2)$ auch der Begriff Mean Square Error (MSE) benutzt. Man kann den mittleren quadratischen Fehler auch als 2-tes Moment bezüglich Θ der Zufallsvariablen $\hat{\Theta}_Y$ auffassen. Ist die Schätzung unverzerrt, so wird aus MSE das 2-te zentrale Moment, die Varianz des Schätzers.

bisher ist die Aufspaltung des Zufallsfehlers nach (2.27) noch nicht berücksichtigt. Die Varianz des Zufallsfehlers Z_Y läßt sich daher weiter zerlegen in

$$\begin{aligned}
 \text{Var}(Z_Y) &= \text{Var}(Z_X + Z_E) \\
 &= \text{Var}(Z_X) + \text{Var}(Z_E) + 2\text{Cov}(Z_X, Z_E).
 \end{aligned}
 \tag{2.33}$$

Neben der Varianz der stichprobenbedingten Zufallsfehler und der Varianz der nicht-stichprobenbedingten Zufallsfehler beeinflusst auch die Abhängigkeit der beiden Zufallsvariablen die Varianz des gesamten Zufallsfehlers Z_Y . Es sind Situationen denkbar, in denen Unabhängigkeit zwischen Z_X und Z_E angenommen werden kann, d. h. $\text{Cov}(Z_X, Z_E) = 0$, aber generell gilt das nicht.

1) Vgl.: Stenger 17, S. 393; Pokropp, 80, S. 21; die Bezeichnung ist jedoch nicht einheitlich. Menges spricht von Quadrat des Gesamtfehlers; vgl.: Menges u. Skala, 73, S. 286.

Unter Berücksichtigung von (2.33) folgt für den MSE

$$\text{MSE} = \text{Var}(Z_X) + \text{Var}(Z_E) + 2 \text{Cov}(Z_X, Z_E) + b_Y^2. \quad (2.34)$$

Stenger, der seine Ausführungen speziell auf Antwortvariabilität bezieht, bezeichnet die einzelnen Komponenten des mittleren quadratischen Fehlers der Schätzung als Stichprobenvarianz, Antwortvarianz, 2 mal Interaktion und Quadrat der Antwortverzerrung. Dabei ist jedoch zu beachten, daß Stenger von der Verzerrung b_X absieht und auch eine andere Abgrenzung zwischen Z_X und Z_E wählt¹⁾. Die Stichprobenvarianz erfaßt bei Stenger neben der Varianz der Schätzung der wahren Merkmalsausprägungen auch den durch die Auswahl bedingten Einfluß der individuellen Verzerrungen auf den Zufallsfehler. Der Begriff Antwortvarianz ist daher enger gefaßt als $\text{Var}(Z_E)$. Die Bezeichnungen können nur für den Spezialfall $E(Y_j) = x_j$, d. h. $b_j = 0$, als identisch benutzt werden.

Kann bei der Erhebung ein individueller Antwortfehler ausgeschlossen werden, so reduziert sich der MSE auf

$$\text{MSE} = \text{Var}(Z_X) + b_X^2. \quad (2.35)$$

Die bisherigen Aussagen über den tatsächlichen Fehler und seine Komponenten sind noch nicht speziell auf einzelne Stichprobenverfahren ausgerichtet. Sie verdeutlichen lediglich die Problematik der Fehlerbehandlung unter Berücksichtigung systematischer Fehler.

1) Vgl.: Stenger, 77, S. 392 ff.

Es werden keinerlei Annahmen über die Art der systematischen Fehler, des zu schätzenden Parameters, über die verwendete Stichprobenfunktion oder die Art der Stichprobenauswahl getroffen, mit der Ausnahme, daß der Auswahl ein Zufallsexperiment zugrunde liegt. All diese Faktoren beeinflussen das konkrete Aussehen der Komponenten des Fehlers T bzw. des MSE.

2.3 Eigenschaften der Fehler in statistischen Untersuchungen

2.3.1 Die Eigenschaften des Zufallsfehlers einer Schätzung

Die bisherigen Ausführungen zu den Fehlern einer statistischen Untersuchung behandeln deren Entstehungsursachen und anschließend wurde der Zufallsfehler und die Verzerrung einer Schätzung näher betrachtet. Damit ist nur wenig über die statistischen Eigenschaften der Fehler ausgesagt.

Die Ausführungen in der Literatur zu den Eigenschaften der Zufallsfehler sind auf den stichprobenbedingten Zufallsfehler Z_X ausgerichtet. Inwieweit auch der nicht-stichprobenbedingte Zufallsfehler Z_E , der ja durch die Antwortfehler in der Erhebung entsteht, erfaßt ist, läßt sich besser im Anschluß an die Behandlung dieser Fehlerquelle in der einfachen bzw. geschichteten Zufallsauswahl überprüfen. In Kapitel 4 werden die Unterschiede in den statistischen Eigenschaften von Z_X und Z_E angesprochen.

Die stichprobenbedingten Zufallsfehler unterscheiden sich in folgenden Eigenschaften¹⁾ von der Verzerrung der Schätzung:

(1) Der Zufallsfehler Z_X entzieht sich der exakten Bestimmung, da i. d. R. der Erwartungswert des Schätzers unbekannt bleibt. Über die Verteilung der Zufallsvariablen $\hat{\theta}_X$, des Schätzers, sind jedoch Wahrscheinlichkeitsaussagen über die Größenordnung des mög-

1) Vgl.: Menges u. Skala, 73, S. 273; Statistisches Bundesamt, 60, S. 49 f.

lichen Zufallsfehlers ableitbar¹⁾). Besondere Bedeutung kommt der Varianz des Schätzers zu, denn mit $\text{Var}(\hat{\theta}_X) = \text{Var}(Z_X)$ ist ein Maß für den Zufallsfehler gegeben; $\text{Var}(Z_X)$ mißt die Wiederholungsgenauigkeit²⁾ oder Präzision³⁾ der Schätzung. Sie kann - und das ist entscheidend - aus der Stichprobe $\underline{X}$, wie sie in (2.10) eingeführt ist, geschätzt werden.

Geben wir eine Wahrscheinlichkeit γ vor, so können wir ein Intervall für den stichprobenbedingten Zufallsfehler Z_X um den Erwartungswert $E(Z_X) = 0$ angeben, in das die Realisationen der Z_X mit der vorgegebenen Wahrscheinlichkeit fallen

$$P(-A_1 \leq Z_X \leq A_2) = \gamma; \quad (2.36)$$

$A_1, A_2 > 0$ ⁴⁾.

Meistens wird die Fehleraussage jedoch nicht auf den Zufallsfehler bezogen, sondern auf den Schätzer $\hat{\theta}_X$ selbst; $\hat{\theta}_X$ ist lediglich eine lineare Transformation von Z_X um $E(\hat{\theta}_X)$. Die Realisationen der Zufallsvariablen $\hat{\theta}_X$ liegen mit vorgegebener Wahrscheinlichkeit γ in einem Intervall um den Erwartungswert $E(\hat{\theta}_X)$:

$$P(E(\hat{\theta}_X) - A_1 \leq \hat{\theta}_X \leq E(\hat{\theta}_X) + A_2) = \gamma. \quad (2.37)$$

Die Intervalllänge, die höchstens zu erwartende Abweichung, hängt von der vorgegebenen Wahrscheinlichkeit γ , dem Verteilungstyp des Schätzers und der Varianz des Schätzers ab.

Die Intervallgrenzen lassen sich bei vorgegebener Wahr-

1) Vgl.: z. B. Bleymüller u. a., 83, S. 85 f.

2) Vgl.: Statistisches Bundesamt, 60, S. 21.

3) Vgl.: Krug u. Nourney, 82, S. 158.

4) Die Intervallgrenzen A_1, A_2 sind vom Verteilungstyp und der Varianz der Zufallsvariablen Z_X bzw. vom Schätzer $\hat{\theta}_X$ abhängig. Im Zusammenhang mit dem speziellen Mittelwertschätzer bei einfacher Zufallsauswahl wird A_1 bzw. A_2 präzisiert (vgl. S. 70 dieser Arbeit).

scheinlichkeit aus der bekannten Verteilung des Schätzers bestimmen, wobei jeweils die Hälfte der vorgegebenen Wahrscheinlichkeit auf die Bereiche $\hat{\theta}_X \leq E(\hat{\theta}_X)$ und $\hat{\theta}_X \geq E(\hat{\theta}_X)$ entfällt.

Dieser direkte Schluß von der Gesamtheit auf den Schätzer kann auch umgekehrt werden in den indirekten Schluß: Bei vorgegebener Wahrscheinlichkeit γ weicht der Erwartungswert des Schätzers höchstens um A_2 nach unten bzw. A_1 nach oben vom Schätzer $\hat{\theta}_X$ ab,

$$P(\hat{\theta}_X - A_2 \leq E(\hat{\theta}_X) \leq \hat{\theta}_X + A_1) = \gamma. \quad (2.38)$$

Das Ergebnis der Schätzung ist ein Zufallsintervall

$$[\hat{\theta}_X - A_2; \hat{\theta}_X + A_1] \quad (2.39)$$

und heißt Konfidenzintervall¹⁾. Die Wahrscheinlichkeit γ wird auch Konfidenzniveau genannt.

In den meisten Fällen genügt der Schätzer einer symmetrischen Verteilung (Normalverteilung, t-Verteilung); damit gilt $A_1 = A_2$, d. h. es entstehen symmetrische Intervalle um den Schätzer. Im Konfidenzintervall ist der stichprobenbedingte Zufallsfehler Z_X berücksichtigt, wobei jedoch auch größere Abweichungen möglich, wenn auch unwahrscheinlich sind.

(2) Betrachtet man die Realisationen der Zufallsvariablen X_i in der Stichprobe $\underline{X}$ und untersucht deren Abweichungen vom Erwartungswert $E(X_i) = \mu$, so sind sowohl positive als auch negative Abweichungen zu erwarten. Durch Aggregation, wie sie im Schätzer $\hat{\theta}_X$ erfolgt, heben sich die Fehler zum Teil auf. Es findet ein Fehlerausgleich bei der Aggregation statt.

1) Vgl.: Menges, 72, S. 285; Bleymüller u. a., 83, S. 86. Die Bezeichnung 'Konfidenzintervall' bezieht sich auch auf ein realisiertes Intervall, d. h., wenn für $\hat{\theta}_X$ ein Schätzwert eingesetzt wird.

(3) Zufallsfehler treten nur in Zufallsstichproben auf, d. h. bei statistischen Untersuchungen, denen ein Zufallsexperiment zugrunde liegt. Totalerhebungen weisen daher keinen stichprobenbedingten Zufallsfehler Z_X auf.

Die drei hier abgeleiteten Eigenschaften von Z_X folgen aus dem Zufallsprozeß, der Entstehungsursache des Fehlers. Um die Unterschiede zwischen stichprobenbedingtem Zufallsfehler und Verzerrung herausarbeiten zu können, ist eine Übertragung bzw. Abgrenzung der drei o. g. Eigenschaften auf bzw. zu der Verzerrung sinnvoll.

2.3.2 Eigenschaften der Verzerrung einer Schätzung

Bei der Betrachtung der Eigenschaften¹⁾ der Verzerrung eines Schätzers spielt es keine Rolle, ob diese durch Antwortfehler oder durch systematische Fehler, die sich global verzerrend auf die Schätzung auswirken, entstehen. Entscheidend ist die Gesamtverzerrung der Schätzung.

(1) Die Gesamtverzerrung b_Y entzieht sich, wie der Zufallsfehler, einer exakten Bestimmung; sowohl $E(\hat{\Theta}_Y)$ als auch Θ sind unbekannt. Im Gegensatz zu dem stichprobenbedingten Zufallsfehler läßt die Stichprobe keine Aussage über b_Y zu. Die traditionelle Fehlerrechnung im Konfidenzintervall berücksichtigt die Verzerrung b_Y nicht.

Die beiden Bestimmungsgrößen der Verzerrung, Θ und $E(\hat{\Theta}_Y)$, sind keine Zufallsvariablen; der bias b_Y ist daher mit der Wahrscheinlichkeitsrechnung nicht erfassbar. Nur zusätzliche Kontrollmaßnahmen geben Auskunft über Größe und Richtung der Verzerrung²⁾.

1) Vgl.: Menges u. Skala, 73, S. 283.

2) Vgl.: Statistisches Bundesamt, 60, S. 110.

(2) Die Verzerrung der Schätzung ist bei gegebenem Erhebungssystem eine Konstante. Wird die Erhebung mehrmals durchgeführt, so ist jeweils der bias gleich. Lediglich bei einem Wechsel des Erhebungssystems ist eine andere Verzerrung zu erwarten.

Eine durch die Aggregation im Schätzer entstehende Ausgleichstendenz der Verzerrung kann nicht unterstellt werden.

Zwar sind Situationen denkbar, in denen die verschiedenen systematischen Fehler gegenläufig sind und sich so zum Teil aufheben, mit zunehmendem Stichprobenumfang sinkt die Verzerrung jedoch nicht, der bias bleibt bestehen.

Zwei Möglichkeiten, wie sich systematischen Fehler gegenseitig aufheben, sind zu unterscheiden:

1. Die aus verschiedenen Verzerrungsursachen entstehenden Komponenten der Gesamtverzerrung weisen unterschiedliche Vorzeichen auf. Beispielsweise kann die Nichtbeantwortung der Fragen durch einzelne ausgewählte Untersuchungseinheiten den Schätzer $\hat{\theta}_Y$ vergrößern, während Verzerrungen der Stichproben durch ungleiche Auswahlwahrscheinlichkeiten oder Mängel der Erfassungsgrundlage die Schätzung verringern. Die resultierende Gesamtverzerrung ist dann geringer als die der einzelnen Verzerrungsursachen.

In diesem Fall ist der Verzerrungsausgleich jedoch nicht durch die Aggregation der individuellen Antworten im Schätzer erfolgt, sondern unabhängig vom Schätzer durch die gegengerichteten Komponenten der Gesamtverzerrung.

2. Bestehen systematische Fehler in den Daten, so können die individuellen Antwortverzerrungen b_j der Einheiten aus G Unterschiede aufweisen. Die Aggregation der Antwortfehler in $E(\hat{\theta}_E)$ führt zu einem gegenseitigen Ausgleich der b_j , wobei jedoch nicht ein vollständiger Ausgleich erfolgen muß. Es bleibt die

resultierende Verzerrung $b_E = E(\hat{\theta}_E)$ bestehen. Auch in diesem Fall ist der Fehlerausgleich unabhängig von der Stichprobenfunktion bzw. vom Stichprobenumfang.

Zusammenfassend bleibt festzuhalten, daß ein möglicherweise bestehender Verzerrungsausgleich von grundlegend anderer Natur ist, wie die Aggregationseigenschaft des stichprobenbedingten Zufallsfehlers. Der Stichprobenumfang hat keinerlei Einfluß auf den Grad des Fehlerausgleichs.

(3) Aus (2) ergibt sich, daß auch Totalerhebungen Verzerrungen aufweisen können, denn mit zunehmendem Stichprobenumfang bleibt der bias erhalten. Das heißt jedoch nicht, daß b_Y bei Stichproben- und Totalerhebung gleich sind, denn mit zunehmendem Stichprobenumfang ist meist ein Wechsel im Erhebungssystem notwendig. Weil bei Stichprobenerhebungen mehr Sorgfalt auf die Untersuchung der ausgewählten Einheiten verwendet werden kann, weisen ihre Ergebnisse oft geringere Verzerrungen auf als die entsprechender Totalerhebungen¹⁾.

Die bisherige Beschreibung der Eigenschaften einer Verzerrung grenzt diese von den Eigenschaften der stichprobenbedingten Zufallsfehler ab. Die wichtigen Eigenschaften der Zufallsfehler - Abschätzbarkeit und Ausgleichstendenz - werden von der Verzerrung nicht erfüllt. Streng genommen sind die Eigenschaften (1) und (2) bei der Verzerrung 'Negativ'-Eigenschaften; die Verzerrung ist durch diese nicht beschrieben, es ist lediglich gesagt, was der bias nicht erfüllt. Die systematischen Fehler und ihre Auswirkungen auf die Schätzaussage sind mit den auf die Zufallsfehler ausgerichteten statistischen Methoden nicht oder nur unzureichend erfaßt.

1) Vgl.: Statistisches Bundesamt, 60, S. 49.

Bestehen systematische Fehler in einer Schätzung, so ist der Erwartungswert des Schätzers $\hat{\theta}_Y$ um den konstanten Bias b_Y verschoben. Man kann dies als lineare Transformation der Verteilung von $\hat{\theta}_Y$ auffassen. Ob auch die Varianz des Schätzers durch die systematischen Fehler beeinflusst wird, bedarf einer genaueren Betrachtung. In Schaubild 2.1 ist die Situation einer verzerrten Schätzung dargestellt.

Die Verzerrung b_Y verändert die Wahrscheinlichkeit, mit der das Konfidenzintervall nach (2.39) den zu schätzenden Parameter θ tatsächlich überdeckt. In Kapitel 3 wird dieser Aspekt genauer betrachtet.

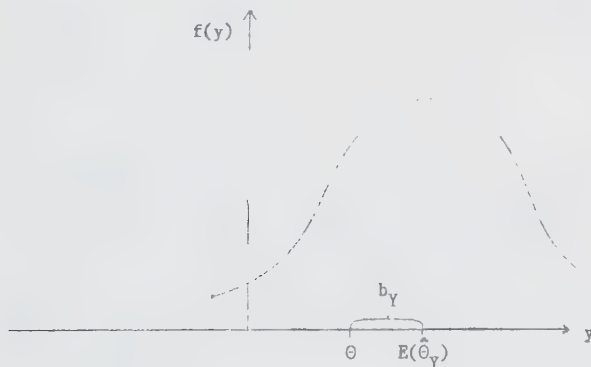


Schaubild 2.1: Darstellung der Verzerrung einer Schätzung; mit y sind die Realisationen des Schätzers $\hat{\theta}_Y$ bezeichnet, $f(y)$ ist die Dichte der Verteilung von $\hat{\theta}_Y$, θ ist der zu schätzende Parameter.

2.4 Kontrollmaßnahmen und ihre Bedeutung für Stichprobenerhebungen

Zusammenfassend bleibt festzuhalten, daß es eine Vielzahl verschiedener Fehlerquellen in statistischen Untersuchungen gibt, die natürlich nicht immer alle gleichzeitig und mit gleicher Intensität auftreten. Je nach dem konkreten Erhebungsgegenstand und den Organisations- und Rahmenbedingungen fallen die Fehler anders aus. Weil die auftretenden Fehler in Stichprobenerhebungen stark variieren, ist das Problem in jedem Einzelfall gesondert zu lösen.

Bereits bei der Planung einer Erhebung sind Maßnahmen zur Fehlerbehandlung mit zu berücksichtigen. Diese Kontrollen können verschiedene Zielsetzungen verfolgen¹⁾:

- Nachträglich wird versucht, durch Kontrollen die Größe der entsprechenden Verzerrung zu ermitteln. Diese deskriptiven Kontrollen sind geeignet, die Güte der Ergebnisse genauer zu kommentieren.
- Die operativen Kontrollen ermöglichen eine Bereinigung der Ergebnisse von Verzerrungen, denn sie sollen die während der Erstellung der Statistik entstehenden Fehler ermitteln.

Korrekt durchgeführte operative Kontrollen ermöglichen ein Reduzieren der Fehler in den Ergebnissen, sie sind deshalb den deskriptiven vorzuziehen. Dem stehen zum Teil prinzipielle, wie beiden Zufallsfehlern, zum Teil auch Wirtschaftlichkeitsüberlegungen entgegen.

1) Vgl.: Statistisches Bundesamt, 60, S. 110 f;
Krug u. Nourney, 82, S. 202 f.

Die Fehlerrechnung zur Quantifizierung der Zufallsfehler könnte zu den deskriptiven Kontrollen gerechnet werden, denn sie kann den tatsächlich realisierten Zufallsfehler nicht angeben¹⁾. Die Varianz des Schätzers ist lediglich ein Maß für die zufällige Abweichung des Schätzers von seinem Erwartungswert.

Im Rahmen der Kontrollmaßnahmen geht es mehr darum, die systematischen Fehler in der Erhebung aufzudecken und, wenn möglich, deren Auswirkungen zu ermitteln. Außerdem sind Möglichkeiten der Vermeidung systematischer Fehler mit zu beachten, einmal weil es besser ist, die Verzerrung der Schätzung nicht erst auftreten zu lassen, als nachträglich die Schätzung zu korrigieren. Zum anderen ist dann auch eine Reduzierung der Zufallsfehler mit systematischem Ursprung möglich.

Es ist bisher noch nicht gelungen, geschlossen die Behandlungsmöglichkeiten der Fehler und die durch diese entstehenden Schätzprobleme zu beschreiben. Die Arbeiten behandeln meist Einzelaspekte²⁾. Nur zum Teil können durch theoretische Überlegungen Aussagen über die Fehlerauswirkungen getroffen werden (z. B. Nichtbeantwortung), meist müssen jedoch geeignete Verfahren gewählt und organisatorische Maßnahmen getroffen werden, um die Fehler zu quantifizieren und ihre Auswirkungen auf die Schätzaussage zu berücksichtigen.

Die Maßnahmen zur Vermeidung von Fehlern sind mit zusätzlichen Kosten verbunden. Es geht deshalb bei der Fehlerbehebung nicht nur darum, mit allen zur Verfü-

1) Üblicherweise beziehen sich Kontrollmaßnahmen aber auf Fehlerquellen, die zu Verzerrungen der Schätzung führen können.

2) Vgl.: z. B. Cochran, 72, S. 422 ff (Nichtbeantwortung); Sukhatme u. Sukhatme, 70, S. 390 ff (Interviewereinfluß); Cochran, 72, S. 437 ff; Stenger, 77, S. 392 ff; Pokropp, 80, S. 201 ff (Antwortvariabilität).

gung stehenden Mitteln die Fehler zu vermeiden, sondern auch um den wirtschaftlichen Einsatz der zur Verfügung stehenden Mittel. Die geforderte Genauigkeit der Schätzung, der tolerierte Fehler, ist in diesem Zusammenhang mit zu berücksichtigen. Die Reduzierung der Fehler darf sich deshalb nicht nur auf die Verzerrung beziehen, denn beim wirtschaftlichen Mitteleinsatz sind auch die Zufallsfehler zu beachten. Letztere sind über das Stichprobenverfahren und den Stichprobenumfang, aber auch durch die Reduzierung der Fehler in den Daten vom Mitteleinsatz abhängig.

Um die Kosten zu reduzieren, erfolgen Kontrollen, oft "auf repräsentativer Basis"¹⁾, d. h. stichprobenartig werden einzelne Erhebungsvorgänge überprüft. In diesen Fällen ist die Zielsetzung nicht mehr die Vermeidung von systematischen Fehlern, sondern Größe und Richtung der bestehenden Fehler sollen ermittelt werden. Ob die konkreten Kontrollen operativen oder deskriptiven Charakter haben, hängt im wesentlichen von der Durchführung der Maßnahmen ab.

In der Vorbereitungsphase einer Erhebung ist eine eingehende Kontrolle der Erhebungs- und Auswahlgrundlage entscheidend für die Vermeidung einer Verzerrung. Ein wichtiger Grund für die Durchführung einer Probeerhebung ist die Kontrolle der Fragebogengestaltung.

Es ist schwierig und nur mit hohem finanziellen Aufwand möglich, Fehler in der Datengewinnung, d. h. bei der direkten Befragung, zu vermeiden oder zu reduzieren.

1) Statistisches Bundesamt, 60, S. 110.

Eine Möglichkeit, bei Befragungsgegenständen mit voraussichtlich starken Antwortfehlern diese zu vermeiden bzw. zu reduzieren, ist das Verfahren des 'randomized response'¹⁾. Es gibt den Antwortenden die Möglichkeit ehrlich zu antworten, ohne für sie unangenehme Fakten oder Meinungen zugeben zu müssen. Allerdings wird durch die Verwendung einer Zufallsauswahl während der Befragung der Zufallsfehler der Schätzung erhöht. Das Verfahren bezieht sich auf qualitative Merkmale.

Die Einflüsse der Antwortvariabilität auf die Varianz der Schätzung läßt sich mit dem Verfahren der 'interpenetrating subsamples'²⁾ oder durch zwei- bzw. mehrmalige Befragung der ausgewählten Elemente³⁾ ermitteln. Dies erfordert, insbesondere bei der Wiederholungsbefragung, zusätzliche finanzielle Mittel.

Die Verzerrung der Schätzung durch Antwortfehler läßt sich nur dann ermitteln, wenn ein 'geeichtetes' Erhebungsverfahren existiert⁴⁾, mit dem die x_j bzw. die b_j der ausgewählten Untersuchungseinheiten ermittelt werden können. Mit diesem Erhebungsverfahren könnte eine Mittelwertschätzung ohne Antwortfehler erfolgen. Existiert ein solches Verfahren nicht, so kann der Ansatz der 'wahren' Merkmalsausprägung der u_j kaum noch aufrechterhalten werden.

Teilweise sind die verzerrungsfreien Erhebungsverfahren mit höheren Kosten verbunden, so daß es wirtschaftlicher ist, bei einer größeren Stichprobe Antwortfehler zuzulassen und anschließend in einer Unterstichprobe geringeren Umfangs mit größerer Sorg-

1) Das Verfahren geht auf S. L. Warner zurück. Vgl.: Warner, 65, S. 63 ff; vgl. auch Pokropp, 80, S. 197 ff.

2) Zum Verfahren vgl.: Stenger, 77, S. 397 ff; Cochran, 72, S. 447 ff; vgl. auch Pokropp, 80, S. 214.

3) Vgl.: Pokropp, 80, S. 214 f.

4) Vgl.: Pokropp, 80, S. 213.

falt, mit besser geschulten Interviewern zu versuchen, verzerrungsfreie Antworten der Elemente zu erhalten. Aus der Differenz der Befragungsergebnisse läßt sich dann die Verzerrung der Gesamtstichprobe schätzen. Einen solchen Ansatz benutzten Strecker und andere zur Ermittlung der Verzerrung bei Erhebungen aus dem Bereich der Agrarstatistik¹⁾.

Kontrollen im Bereich der Aufbereitung können auf vorhandenem Datenmaterial aufbauen. Sie sind deshalb relativ leicht zentral durchzuführen. Die Ergebnisse können oft um die festgestellten Verzerrungen bereinigt werden. Damit haben die Aufbereitungskontrollen operativen Charakter²⁾.

Ein wesentlicher Vorteil von Kontrollmaßnahmen ist es, Erkenntnisse über wichtige Fehlerquellen in praktischen Erhebungen zu gewinnen. Diese können dann in die Planung und Durchführung späterer Erhebungen eingehen³⁾.

Dieser Aspekt ist vor allem deshalb wichtig, weil (fast) alle Maßnahmen zur Vermeidung und zur Bestimmung von systematischen Fehlern bereits bei der Planung mit berücksichtigt werden müssen. Nachträglich ist es kaum möglich die Verzerrung zu bestimmen, wenn nicht bereits bei der Planung der Untersuchung diese Maßnahmen mit vorgesehen sind. Dies ist erstens wegen der Verteilung der Kosten auf die einzelnen Erhebungsphasen der Fall, aber auch zum Teil aus prinzipiellen Gründen.

Im Zusammenhang zwischen Fehler, Fehlergröße und der Bedeutung für die Schätzung ist auch die Fehlerfortpflanzung zu erwähnen. Sollen die ermittelten Schätz-

1) Vgl.: Strecker u. a., 83, S. 24 ff; Strecker, 78, S. 100 ff.

2) Vgl.: Statistisches Bundesamt, 60, S. 111.

3) Vgl.: Statistisches Bundesamt, 60, S. 110.

werte zur Berechnung von Maßzahlen herangezogen werden, so ist bei der Bestimmung der notwendigen Genauigkeit zu beachten, wie sich der Fehler auf die zu berechnende Maßzahl auswirkt. Speziell bei Quotienten ist besondere Aufmerksamkeit angebracht¹⁾.

Die operativen Kontrollmaßnahmen, die in der Lage sind, die Verzerrung zu ermitteln, sind fehlerspezifisch, d. h., wenn eine Verzerrung ermittelt worden ist, so bezieht sich diese nur auf die betrachtete(n) Fehlerquelle(n). Es können dann trotzdem weitere Verzerrungen bestehen. Wird beispielsweise durch eine Unterstichprobe die Verzerrung durch die Antwortfehler geschätzt, bleibt eine möglicherweise bestehende Verzerrung durch Fehler in der Auswahlgrundlage weiter bestehen.

Darzustellen, welche Kontrollmaßnahmen in der Praxis möglich und wie diese durchzuführen sind, ist nicht Ziel dieser Arbeit. Der Schwerpunkt der Betrachtungen liegt auf der "stochastischen Behandlung"²⁾ der Fehler; die Auswirkungen der nicht-stichprobenbedingten Fehler auf die Schätzaussage werden näher untersucht.

1) Zu den Fehlerfortpflanzungsregeln vgl.: Anderson, 65, S. 80 ff; Kallmeyer, 56, S. 21; Anderson, 54, S. 41 ff.

2) Menges u. Skala, 73, S. 283.

3 Auswirkungen systematischer Fehler auf Schätzungen bei korrekten individuellen Antworten

3.1 Grundlagen

In diesem Kapitel soll untersucht werden, wie sich die konstante Verzerrung b_x auf die Ergebnisse der einfachen Schichtenauswahl auswirkt. Wir setzen dabei voraus, daß die Merkmalsträger u_j aus G mit ihren wahren Merkmalsausprägungen x_j antworten. Die systematischen Fehler in den Antworten bleiben daher vorerst unberücksichtigt.

Beim Schichtenauswahlverfahren gilt für jede Schicht, isoliert betrachtet, das Modell der einfachen Zufallsauswahl, und es ist daher sinnvoll, mit diesem grundlegenden Stichprobenverfahren zu beginnen. Zusätzlich ergibt sich dann die Möglichkeit, die Auswirkungen der Verzerrungen bei der einfachen Zufallsauswahl mit denen bei einfacher Schichtenauswahl zu vergleichen.

Das Modell der einfachen Zufallsauswahl ist in der Literatur umfassend beschrieben¹⁾. In dieser Arbeit genügt daher eine knappe Beschreibung der Zusammenhänge, die zum Verständnis der Auswirkungen einer Verzerrung auf die Ergebnisse der Schätzung notwendig sind. Auf Beweise und die Diskussion der Eigenschaften der Schätzer wird weitgehend verzichtet.

In den Untersuchungen will ich mich auf quantitative Merkmale beschränken. Die so gewonnenen Ergebnisse

1) Vgl.: z. B. Pokropp, 80, S. 26 ff; Stenger, 71, S. 66 ff; Cochran 72, S. 34 ff, Hansen u. a., I, 53, S. 110 ff.

lassen sich prinzipiell auch auf qualitative Merkmale übertragen. Um die Anzahl der verwendeten Symbole nicht unnötig groß werden zu lassen und so die Übersichtlichkeit der Arbeit zu beeinträchtigen, erscheint eine Beschränkung auf das arithmetische Mittel als zu schätzender Parameter angebracht.

3.2 Die Verzerrung bei einfacher Zufallsauswahl

3.2.1 Das Modell der einfachen Zufallsauswahl

3.2.1.1 Die Stichprobe als Zufallsvektor

Gegeben ist die Grundgesamtheit G als Menge der N für die Untersuchung relevanten Merkmalsträger u_j , $j = 1, 2, \dots, N$. Das Untersuchungsmerkmal, hier sei ein quantitatives Merkmal unterstellt, ordnet über das Antwortverhalten (2.6) jedem u_j aus G eine reelle Zahl x_j als Merkmalsausprägung zu. Aus der Menge G entsteht dann die Folge der Merkmalsausprägungen

$$x = (x_j / j = 1, 2, \dots, N). \quad (3.1)$$

Auf G_x ist das arithmetische Mittel μ nach (2.7) als zu schätzender Parameter definiert, d. h. speziell gilt $\Theta = \mu$. Als Maß für die Streuung der Merkmalsausprägungen in G_x ist

$$\sigma^2 = \frac{1}{N} \sum_{j=1}^N (x_j - \mu)^2, \quad (3.2)$$

die mittlere quadratische Abweichung der Merkmalsausprägungen vom arithmetischen Mittel, geeignet. Man bezeichnet σ^2 auch kurz als Varianz der Grundgesamtheit. Häufig wird auch

$$s^2 = \frac{1}{N-1} \sum_{j=1}^N (x_j - \mu)^2 = \frac{N}{N-1} \sigma^2 \quad (3.2*)$$

als Maß für die Streuung benutzt. Pokropp und Stenger bezeichnen s^2 auch als korrigierte Varianz¹⁾.

Der Parameter μ soll mit einer Zufallsstichprobe vom Umfang n geschätzt werden. Statt der Stichprobe $\underline{U}$ aus G können wir, da wir das Antwortverhalten (2.6) unterstellen, sofort entsprechend (2.10)

$$\underline{X} = (X_i / i = 1, 2, \dots, n)$$

aus G_x betrachten. Der Index i identifiziert die nacheinander auszuführenden Auswahlvorgänge. Die möglichen Ergebnisse der Zufallsauswahl i sind die in G_x zusammengefaßten Merkmalsausprägungen, d. h. G_x ist die Folge der Realisationen der Zufallsvariablen X_i .

Die einfache Zufallsauswahl impliziert, daß alle zu G gehörenden Elemente bei der Auswahl der Stichprobe gleichberechtigt sind, unabhängig vom Auswahlmodus²⁾. Wird nach jedem Auswahlvorgang das gezogene Element weiter mit berücksichtigt, d. h. wird mit Zurücklegen gezogen, ist die Wahrscheinlichkeit, daß beim i -ten Auswahlvorgang das Element x_j gezogen wird, für alle Elemente gleich

$$P(X_i = x_j) = \frac{1}{N}, \quad i = 1, 2, \dots, n, \\ j = 1, 2, \dots, N. \quad (3.3)$$

Man kann dies als klassische Wahrscheinlichkeit auffassen, so daß mit (3.3) für alle N Elemente aus G

1) Vgl.: Stenger, 71, S. 67; Pokropp, 80, S. 9; durch die Verwendung von (3.2*) werden die Formeln für die Varianz des Schätzers teilweise einfacher. Im folgenden werden die Summationsgrenzen, sofern keine Verwechslungen zu befürchten sind, weggelassen.

2) Vgl.: Pokropp, 80, S. 27 und S. 29.

das Wahrscheinlichkeitsmaß für X_i gegeben ist. Der klassische Wahrscheinlichkeitsbegriff erfüllt die Bedingungen, die an ein Wahrscheinlichkeitsmaß gestellt sind.

Beim Ziehen mit Zurücklegen sind die X_i aus $\underline{X}$ unabhängige Zufallsvariablen, denn durch das Zurücklegen der gezogenen Einheiten haben die bereits vorgenommenen Auswahlvorgänge keinen Einfluß auf den Zufallsprozeß i .

Für $\underline{X}$ ergeben sich insgesamt N^n Realisationen, die alle mit der gleichen Wahrscheinlichkeit $(N^n)^{-1}$ eintreten¹⁾.

Beim Ziehen mit Zurücklegen ist es möglich, daß einzelne Elemente aus G mehrmals in die Stichprobe gelangen. Das mehrmalige Auftreten einzelner Elemente bringt jedoch keine zusätzlichen Informationen. Dies vermeidet der Auswahlmodus Ziehen ohne Zurücklegen. Die Zufallsvariablen X_i , $i = 1, 2, \dots, n$, haben alle die gleiche Verteilung, d. h. die Menge der Realisationen ist jeweils G_X und die Wahrscheinlichkeit, daß X_i eine bestimmte Realisation annimmt, ist²⁾

$$\begin{aligned} P(X_i = x_j) &= P(X_1 \neq x_j, \dots, X_{i-1} \neq x_j, X_i = x_j, \dots) \\ &= \frac{1}{N}, \quad i = 1, 2, \dots, n. \end{aligned} \quad (3.4)$$

Die Zufallsvariablen in $\underline{X}$ sind allerdings nicht mehr unabhängig: Die Kovarianz zwischen zwei beliebigen der n Zufallsvariablen beträgt³⁾

1) Die Menge der Elementarereignisse von $\underline{X}$, mit G^n bezeichnet, entsteht als Variation mit Wiederholung von n Elementen aus G . Die Anzahl der Elemente aus G^n ist damit N^n . Die Permutationen der n Elemente der Stichprobe sind dabei als verschiedene Realisationen von $\underline{X}$ angenommen.

2) Zum Beweis vgl.: Pokropp, 80, S. 30 f.

3) Vgl.: Pokropp, 80, S. 30.

$$\text{Cov}(X_k, X_i) = -\frac{1}{N-1} \sigma^2, \quad k, i = 1, 2, \dots, n, \\ k \neq i. \quad (3.5)$$

Die Zahl der möglichen Realisationen von $\underline{X}$ reduziert sich durch das Ziehen ohne Zurücklegen von N^n auf¹⁾

$$n! \binom{N}{n} = N(N-1) \dots (N-n+1).$$

Generell kann nicht entschieden werden, welcher Auswahlmodus vorzuziehen ist. Beim Ziehen mit Zurücklegen entsteht wegen der möglichen wiederholten Zählung einzelner Elemente ein Informationsverlust. Dieser ist dann besonders groß, wenn der Stichprobenumfang, relativ zur Grundgesamtheit, hoch ist. Die Beweisführung und die bei der Schätzung zu verwendenden Formeln sind im Fall mit Zurücklegen wegen der Unabhängigkeit der X_i in $\underline{X}$ jedoch einfacher. Ist der Stichprobenumfang gegenüber N sehr klein²⁾, können die Formeln der Auswahl mit Zurücklegen als Approximation für die einfache Zufallswahl ohne Zurücklegen benutzt werden.

Bevor wir uns mit dem Schätzer selbst beschäftigen, ist es sinnvoll, kurz die Parameter der Zufallsvariablen X_i anzugeben: Es gilt für alle $i = 1, 2, \dots, n$

$$E(X_i) = \sum_j \frac{1}{N} x_j = \mu; \quad (3.6)$$

$$\text{Var}(X_i) = E(X_i - E(X_i))^2 \\ = \sum_j \frac{1}{N} (x_j - \mu)^2 = \sigma^2. \quad (3.7)$$

1) Einige Autoren betrachten die Stichprobe bei einfacher Zufallsauswahl ohne Zurücklegen nicht als Zufallsvektor, sondern als Teilmenge aus G . Diese ist eine eindimensionale Zufallsvariable mit $\binom{N}{n}$ Realisationen (Kombination ohne Wiederholung mit n aus N). Die $n!$ Permutationen der in den Realisationen enthaltenen n Elemente sind dabei nicht unterschieden. Jede der Realisationen hat die gleiche Wahrscheinlichkeit $\binom{N}{n}^{-1}$ gezogen zu werden (vgl.: Cochran, 72, S. 34 f; Hansen u. a., II, S. 20 f). Die Ergebnisse für die Schätzung sind für beide Verfahren natürlich gleich.

2) z. B. für $\frac{n}{N} < 0,05$; vgl.: Cochran, 72, S. 40.

3.2.1.2 Der Mittelwertschätzer bei einfacher Zufallsauswahl

Als Schätzer für μ der Gesamtheit G ist das arithmetische Mittel der Zufallsvariablen in $\underline{X}$,

$$\bar{X} = \frac{1}{n} \sum_i X_i, \quad (3.8)$$

geeignet. Als Erwartungswert der Zufallsvariablen $\bar{X}$ ergibt sich mit (3.6)

$$E(\bar{X}) = E\left(\frac{1}{n} \sum_i X_i\right) = \mu, \quad (3.9)$$

so daß mit $\bar{X}$ ein erwartungstreuer Schätzer für μ gegeben ist¹⁾. Dies gilt jedoch nur, wenn keine systematischen Fehler in der Erhebung auftreten.

Um die Präzision der Schätzung bewerten zu können, sind Aussagen über die Varianz des Schätzers notwendig. In die Ableitung von $\text{Var}(\bar{X})$ geht ein, ob die X_i unabhängig sind oder nicht, d. h. es ist eine Fallunterscheidung zwischen den Auswahlmodi Ziehen mit und ohne Zurücklegen notwendig.

Für das Ziehen mit Zurücklegen (mZ) gilt wegen der Unabhängigkeit der Auswahlvorgänge mit (3.7)

$$\text{Var}(\bar{X}_{mZ}) = \text{Var}\left(\frac{1}{n} \sum_i X_i\right) = \frac{\sigma^2}{n}. \quad (3.10)$$

Die Varianz des Schätzers hängt damit von der Varianz σ^2 der Merkmalsausprägungen in G_x und dem Stichprobenumfang n , nicht aber vom Umfang der Gesamtheit ab.

Beim Auswahlmodus Ziehen ohne Zurücklegen ist bei der Berechnung von $\text{Var}(\bar{X})$ die Abhängigkeit der X_i

1) Auf die Überprüfung anderer günstiger Eigenschaften des Schätzers $\bar{X}$ soll hier verzichtet werden.

im Zufallsvektor $\underline{X}$ durch (3.5) zu berücksichtigen.
Es ergibt sich¹⁾

$$\text{Var}(\bar{X}_{0Z}) = \frac{1}{n} \sigma^2 \left(1 - \frac{n-1}{N-1}\right) = \frac{1}{n} \sigma^2 \frac{N-n}{N-1} \quad (3.11)$$

oder mit der korrigierten Varianz s^2 nach (3.2*)

$$\text{Var}(\bar{X}_{0Z}) = \frac{1}{n} s^2 \left(1 - \frac{n}{N}\right). \quad (3.12)$$

Das Verhältnis $\frac{n}{N}$ heißt auch Auswahlssatz der Stichprobe und wird kurz mit f bezeichnet; damit gilt

$$\text{Var}(\bar{X}_{0Z}) = \frac{1}{n} s^2 (1-f). \quad (3.13)$$

Für die Faktoren $\frac{N-n}{N-1}$ und $(1-f)$ (bzw. deren Quadratwurzel, wenn man die Standardabweichung von $\bar{X}$ betrachtet) hat sich in der Literatur die Bezeichnung 'Endlichkeitskorrektur' eingebürgert, denn für unendliche Gesamtheiten gilt $f=0$. Die Bezeichnung ist jedoch nicht glücklich, denn eine Korrektur im eigentlichen Sinn wird mit den Faktoren nicht vorgenommen²⁾. Außerdem fallen die Faktoren bei Stichproben, die mit Zurücklegen gezogen sind, ebenfalls weg.

Der Auswahlmodus Ziehen ohne Zurücklegen liefert eine bessere Schätzung, d. h. die Varianz des Schätzers ist geringer, denn der Endlichkeitskorrekturfaktor ist kleiner eins. Gleichzeitig ist auch die Wirkung einer Vergrößerung des Stichprobenumfanges stärker, denn der Korrekturfaktor sinkt mit zunehmendem n .

Es ist sinnvoll, für die beiden Auswahlmodi eine gemeinsame Darstellung zu definieren³⁾. Man setzt

1) Zum Beweis vgl.: Pokropp, 80, S. 30 f.

2) Vgl.: Menges u. Skala, 73, S. 277; Cochran, 72, S. 40.

3) Vgl.: Pokropp, 80, S. 39 f.

$$\text{Var}(\bar{X}) = \frac{1}{n} \sigma^2 (1 - a \cdot f^*) \quad (3.14)$$

mit $f^* = \frac{n-1}{N-1}$,

wobei beim Auswahlmodus Ziehen mit Zurücklegen $a = 0$ und für das Ziehen ohne Zurücklegen $a = 1$ zu setzen ist.

Der Zusammenhang zwischen der Varianz des Schätzers und der Varianz des Merkmals in der Grundgesamtheit gemäß (3.14) genügt meistens nicht zur Beurteilung einer Schätzung, denn σ^2 bzw. s^2 sind in der Regel unbekannt; sie müssen aus der Stichprobe geschätzt werden.

Die korrigierte Stichprobenvarianz

$$S^2 = \frac{1}{n-1} \sum_i (X_i - \bar{X})^2 \quad (3.15)$$

eignet sich als Schätzer für σ^2 bzw. s^2 . Es gilt speziell, wobei zwischen den Auswahlmodi Ziehen mit bzw. ohne Zurücklegen (mZ bzw. oZ) unterschieden werden muß¹⁾,

$$E(S^2) = \sigma^2 \left(1 + \frac{a}{N-1}\right) = \begin{cases} \sigma^2 & \text{f. mZ} \\ s^2 & \text{f. oZ} \end{cases}$$

Wegen (3.10) bzw. (3.13) ist

$$\widehat{\text{Var}}(\bar{X}) = \frac{1}{n} S^2 (1 - a) \quad (3.16)$$

ein erwartungstreuer Schätzer für $\text{Var}(\bar{X})$ ²⁾.

1) Vgl.: z. B. Pokropp, 80, S. 28 ff.

2) Streng genommen müßte noch die Varianz des Varianzschätzers untersucht werden (vgl. dazu: Cochran, 72, S. 62 f), aber in der Regel unterbleibt dies.

3.2.1.3 Der Zufallsfehler des Schätzers

In Kapitel 2 ist der Zufallsfehler $Z_{\bar{X}}$ einer Schätzung ohne Antwortfehler mit (2.19) definiert als Differenz zwischen Schätzer und seinem Erwartungswert. Für die Mittelwertschätzung bei einfacher Zufallsauswahl gilt speziell

$$Z_{\bar{X}} = \bar{X} - E(\bar{X}). \quad (3.17)$$

Die Zufallsvariable $Z_{\bar{X}}$ hat den Erwartungswert $E(Z_{\bar{X}})=0$ und die Varianz ist gleich der des Schätzers $\bar{X}$, also

$$\text{Var}(Z_{\bar{X}}) = \text{Var}(\bar{X}). \quad (3.18)$$

Bezieht man das bisher für den Schätzer $\bar{X}$ abgeleitete auf den Zufallsfehler, so lassen sich die in Kapitel 2 formulierten Eigenschaften des Zufallsfehlers $Z_{\bar{X}}$ präzisieren.

Mit (3.16) existiert ein Schätzer für die Varianz des Zufallsfehlers $Z_{\bar{X}}$, die Abschätzbarkeit aus der Stichprobe ist damit gegeben.

Die Aggregationseigenschaft von $Z_{\bar{X}}$ kommt im Faktor $\frac{1}{n}$ in (3.14) zum Ausdruck. Mit zunehmendem Stichprobenumfang sinkt der Zufallsfehler, allerdings unterproportional.

Der Zufallsfehler ist beim Auswahlmodus "Ziehen ohne Zurücklegen" kleiner und auch die Steigerung des Stichprobenumfanges wirkt sich stärker aus. Letzteres kommt durch eine Senkung des Endlichkeitskorrekturfaktors mit steigendem n zum Ausdruck. Geht die Stichprobe in eine Totalerhebung über, so gilt $n = N$ und der Endlichkeitsfaktor nimmt den Wert Null an. Die Schätzung weist keinen Zufallsfehler mehr auf, denn es gilt $\text{Var}(Z_{\bar{X}}) = 0$.

3.2.2 Die Beschreibung systematischer Fehler im Modell der einfachen Zufallsauswahl

3.2.2.1 Die Entstehung der Verzerrung

Die bisherigen Betrachtungen zum Modell der einfachen Zufallsauswahl basieren auf der Annahme, daß die Stichprobe $\underline{X}$ nicht durch systematische Fehler verzerrt sei. Im folgenden ist eine mit systematischen Fehlern behaftete Stichprobe unterstellt, die, da wir weiterhin das Antwortverhalten (2.6) voraussetzen, zu einem um

$$b_{\bar{X}} = E(\bar{X}) - \mu \quad (3.19)$$

verzerrten Schätzer $\bar{X}$ führen. Dieser ist nicht mehr erwartungstreu bezüglich μ , denn die Verteilung von $\bar{X}$ ist um $b_{\bar{X}}$ verschoben. Für den Erwartungswert gilt:

$$E(\bar{X}) = \mu + b_{\bar{X}}.$$

Bevor die Einflüsse der systematischen Fehler auf die Schätzaussage besprochen werden sollen, erscheint es sinnvoll, einige Fehlerursachen bei richtigen Antworten der Untersuchungseinheiten im Modell der einfachen Zufallsauswahl detaillierter zu beschreiben.

Wenn man von den Fehlern während der Aufbereitungsphase absieht, sind unter der Voraussetzung korrekter individueller Antworten die Erfassungsfehler und die Stichprobenverfälschungen die wichtigsten Fehlerquellen.

Die Nichtbeantwortung führt nur dann zu einer Verzerrung, wenn der Mittelwert der Merkmalsausprägungen der antwortenden Untersuchungseinheiten von dem der Antwortverweigerer abweicht. Auf eine genaue Ableitung der Verzerrung soll hier jedoch verzichtet werden¹⁾.

1) Die Verzerrung der Schätzung durch Nichtbeantwortung beschreibt z. B. Cochran, 72, S. 415 ff.

In (3.9) ist gezeigt, daß $\bar{X}$ bei einfacher Zufallsauswahl ein erwartungstreuer Schätzer für den wahren Parameter μ der Gesamtheit G ist, d. h. die Fehlerquelle 'Verzerrung der Schätzfunktion' tritt bei diesem Schätzverfahren¹⁾ nicht auf.

Da die Auswahlgrundlage Mängel in der Erhebungsgrundlage mit erfaßt, kann sich die detaillierte Betrachtung der systematischen Fehler bei einfacher Zufallsauswahl auf die Fehler der Auswahlgrundlage und im Auswahlprozeß beschränken.

3.2.2.2 Fehlerhafte Auswahlgrundlage

Die systematischen Fehler in der Auswahlgrundlage bewirken, daß diese nicht mehr die Zielgesamtheit G , sondern eine davon abweichende Gesamtheit

$$G' = \{u'_j / j = 1, 2, \dots, N'\}$$

auflistet. Für diese verzerrte Gesamtheit vom Umfang N' existiert, da wir auch weiterhin das Antwortverhalten (2.6) unterstellen, analog zur wahren Gesamtheit ein Mittelwert

$$\mu' = \frac{1}{N'} \sum_j x'_j$$

mit $x'_j = X(u'_j)$

und eine Varianz

$$\sigma'^2 = \frac{1}{N'} \sum_j (x'_j - \mu')^2,$$

wobei die Symbole entsprechend der wahren Gesamtheit gewählt sind. Der Zusatz "'" kennzeichnet sie als aus der verzerrten Gesamtheit stammend.

1) ebenso beim Schichtenauswahlverfahren, vgl. Kapitel 3 dieser Arbeit.

Die Schätzung $\bar{X}'$ aus G' ist erwartungstreu bezüglich μ'

$$E(\bar{X}') = \mu',$$

sofern keine zusätzlichen systematischen Fehler bestehen. Bezogen auf den wahren Parameter μ ist die Schätzung jedoch um

$$\begin{aligned} \delta \bar{X}' &= E(\bar{X}') - \mu \\ &= \mu' - \mu \end{aligned}$$

verzerrt.

Da die Aussagen der einfachen Zufallsauswahl für die Schätzung aus G' erhalten bleiben, gilt für die Varianz des Schätzers

$$\begin{aligned} \text{Var}(\bar{X}') &= \frac{1}{n} \sigma'^2 (1 - f'^a) \\ \text{mit } f'^a &= \frac{n-1}{N'-1}. \end{aligned}$$

Sie hängt damit nicht mehr von der Varianz des Merkmals X in der wahren Grundgesamtheit G ab.

Wichtig für die Anwendung ist, daß trotz der systematischen Fehler die Varianz des Schätzers aus der Stichprobe geschätzt werden kann:

$$\begin{aligned} \widehat{\text{Var}}(\bar{X}') &= \frac{1}{n} S'^2 (1 - f'^a) \\ \text{mit } f' &= \frac{n}{N}, \end{aligned} \tag{3.20}$$

Mit (3.20) ist gleichzeitig die Varianz des Zufallsfehlers $Z_{\bar{X}'}$ bei verzerrter Auswahlgrundlage geschätzt.

In den Fällen, in denen die Varianz des Merkmals in G von Interesse ist, muß die Verzerrung von S'^2 als Schätzer für σ^2 bzw. s^2 beachtet werden.

3.2.2.3 Systematische Fehler im Auswahlprozeß

Ist der Auswahlmodus mit einem systematischen Fehler behaftet, werden z. B. einige Untersuchungseinheiten bevorzugt, so sind die Wahrscheinlichkeiten für das Ziehen einzelner Elemente nicht mehr gleich. In die Berechnung des Erwartungswertes von $\bar{X}$ gehen die Wahrscheinlichkeiten

$$P(\bar{X}_i = x_j) = p'_j, j = 1, 2, \dots, N,$$

$$\text{mit } \sum_j p'_j = 1$$

ein, wobei für mindestens ein j gilt

$$p'_j \neq \frac{1}{N}.$$

Für $L(\bar{X})$ folgt damit

$$\begin{aligned} L(\bar{X}) &= E\left(\frac{1}{n} \sum_{i=1}^n \bar{X}_i\right) \\ &= \frac{1}{n} \sum_{i=1}^n E(\bar{X}_i) \\ &= \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^N x_j p'_j \\ &= \sum_{j=1}^N x_j p'_j = \mu'. \end{aligned}$$

Wegen des systematischen Fehlers im Auswahlvorgang entsteht die Verzerrung des Schätzers

$$\begin{aligned} b_{\bar{X}} &= L(\bar{X}) - \mu \\ &= \mu' - \mu. \end{aligned}$$

Analog verzerrt der Fehler im Auswahlmodus auch die Varianz des Schätzers, denn bei der Erwartungswertbildung über alle Stichprobenmittelwerte findet die

verzerrte Auswahlwahrscheinlichkeit p'_j der Untersuchungseinheiten Eingang; die Wahrscheinlichkeitsverteilung der Stichproben ist keine Gleichverteilung mehr. Die Varianz des Schätzers kann nicht mehr mit (3.14) aus der Varianz der Grundgesamtheit berechnet werden.

Sind die p'_j für alle Elemente aus G bekannt (was selten der Fall seindürfte), so kann, zumindest beim Auswahlmodus Ziehen mit Zurücklegen, die Verzerrung des Schätzers nachträglich verhindert werden durch die Verwendung des Schätzers bei gegebenen Auswahlwahrscheinlichkeiten¹⁾.

3.2.3 Der tatsächliche Fehler bei einfacher Zufallsauswahl

Der tatsächliche Fehler T erfaßt sowohl den Zufallsfehler $Z_{\bar{X}}$ als auch die Verzerrung der Schätzung $b_{\bar{X}}^2$). Mit dem Mean Square Error (MSE) ist in Abschnitt 2.2.4 ein Maß für T abgeleitet. Für den Schätzer $\bar{X}$ bei einfacher Zufallsauswahl können wir mit (3.16) und (3.14) pruzisieren:

$$MSE_{\bar{X}} = \text{Var}(\bar{X}) + b_{\bar{X}}^2 \quad (3.21)$$

$$= \frac{1}{n} \sigma^2 (1-f) + b_{\bar{X}}^2.$$

Im MSE sind die Abweichungen des Schätzers vom Mittelwert μ quadriert. Um wieder auf die gleichen Linheiten wie bei μ zu kommen, ist es sinnvoll, auch

1) Zum Verfahren vgl.: Pokropp, 80, S. 73 ff.

2) Im folgenden ist eine Unterscheidung zwischen den Verzerrungen aus den verschiedenen Ursachen nicht notwendig. $b_{\bar{X}}$ ist als Gesamtverzerrung der Schätzung bei korrekten individuellen Antworten zu interpretieren.

$$\sqrt{\text{MSE}_{\bar{X}}} =: \text{Gf} \quad (3.22)$$

zu betrachten, für die Menges u. Skala¹⁾ die Bezeichnung Gesamtfehler benutzen. Die Gleichung (3.21) kann nach dem Satz des Pythagoras über rechtwinklige Dreiecke durch das Fehlerdreieck²⁾ (Schaubild 3.1), dargestellt werden.

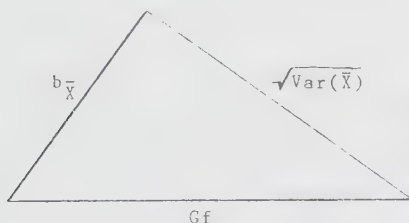


Schaubild 3.1: Das Fehlerdreieck

Die Hypotenuse des rechtwinkligen Dreiecks, die den Gesamtfehler symbolisiert, wird durch die beiden Katheten des Dreiecks, d. h. die Standardabweichung $\sqrt{\text{Var}(\bar{X})}$ des Schätzers und die Verzerrung $b_{\bar{X}}$, bestimmt.

Für den Gesamtfehler läßt sich aus Gleichung (3.21) folgender Ausdruck ableiten:

$$\text{Gf} = \sqrt{\text{Var}(\bar{X}) \left(1 + \frac{b_{\bar{X}}^2}{\text{Var}(\bar{X})} \right)}, \quad (3.23)$$

1) Vgl.: Menges u. Skala, 73, S. 286.

2) Vgl.: Menges und Skala, 73, S. 286; Statistisches Bundesamt, 60, S. 112 f.

wobei der Faktor

$$c := \sqrt{1 + \frac{b^2 \bar{X}}{\text{Var}(\bar{X})}} \quad (3.24)$$

angibt, um wieviel der Gesamtfehler durch die Verzerrung vergrößert wird. Zu beachten ist, daß das Verhältnis aus Quadrat der Verzerrung und Varianz des Schätzers entscheidend für den Einfluß der Verzerrung auf den Gesamtfehler ist. Um die Schreibweise zu vereinfachen, ist es sinnvoll, die relative Verzerrung

$$b^* := \frac{b \bar{X}}{\sqrt{\text{Var}(\bar{X})}} \quad (3.25)$$

einzuführen. Tabelle 3.1 zeigt, wie sich c in Abhängigkeit von b^* verändert.

b^*	c	b^*	c
0,1	1,005	0,6	1,166
0,2	1,020	0,8	1,281
0,3	1,044	1,0	1,414
0,4	1,077	1,5	1,803
0,5	1,118	2,0	2,236

Tabelle 3.1: Einfluß von b^* auf den Gesamtfehler¹⁾

Beträgt die Verzerrung beispielsweise 20 % von $\sqrt{\text{Var}(\bar{X})}$, d. h. $b^* = 0,2$, so steigt der Gesamtfehler lediglich um rund 2 %. Bei einer relativen Verzerrung von $b^* = 0,5$ beträgt die Steigerung bereits rund 12 %.

1) Vgl.: Statistisches Bundesamt, 60, S. 113; die Tabelle vergleichbaren Inhalts ist jedoch anders aufgebaut. Da b^* in c quadratisch eingeht, genügt es, positive b^* zu betrachten.

Aus der Tabelle wird deutlich, daß es wenig Sinn hat, durch eine Vergrößerung des Stichprobenumfanges $\text{Var}(\bar{X})$ immer weiter zu senken, wenn eine Verzerrung vorliegt. Geringfügige Verzerrungen beeinflussen den Gesamtfehler kaum und können unberücksichtigt bleiben. Menges und Skala schlagen vor, die Verzerrungen mit $b^* < 0,25$ zu vernachlässigen, weil dann der Einfluß von $b_{\bar{X}}$ auf den Gesamtfehler noch sehr gering ist¹⁾.

3.2.4 Auswirkungen der systematischen Fehler auf die Schätzaussage bei einfacher Zufallsauswahl

3.2.4.1 Die Konfidenzaussage bei verzerrter Schätzung

Eine mit einem systematischen Fehler behaftete Schätzung ist nicht mehr erwartungstreu bezüglich des wahren Parameters μ . Damit ist bereits ausgesagt, daß die Punktschätzung $\bar{X}$ als Schätzer für μ unzutreffend ist. In der Praxis wird meistens die Intervallschätzung benutzt, d. h. es wird ein Konfidenzintervall mit vorgegebener Irrtumswahrscheinlichkeit für den wahren Parameter formuliert. In diesem Konfidenzintervall ist über die Varianz des Schätzers der Zufallsfehler der Schätzung berücksichtigt.

Welche Auswirkung hat die Verzerrung auf die Wahrscheinlichkeitsaussage des Konfidenzintervalls? Um diese Frage zu beantworten, ist eine genauere Betrachtung der Intervallschätzung notwendig.

1) Vgl.: Menges u. Skala, 73, S. 288.

Will man ein Konfidenzintervall formulieren, so muß der Verteilungstyp des Schätzers bekannt sein. Nach dem zentralen Grenzwertsatz strebt die Summe von n unabhängigen, identisch verteilten Zufallsvariablen mit wachsendem n gegen eine Normalverteilung. Der Schätzer $\bar{X}$ ist daher bei genügend großem Stichprobenumfang asymptotisch normalverteilt. Trotz der Abhängigkeit der X_i beim Auswahlmodus Ziehen ohne Zurücklegen gilt das ebenfalls, sofern der Umfang der Gesamtheit groß genug ist. Als Faustregel wird $N > 2n$ genannt¹⁾. Die Parameter der Verteilung von $\bar{X}$ sind bereits in Abschnitt 3.2 abgeleitet.

Die im Zusammenhang mit den Eigenschaften der Zufallsfehler bereits erwähnte Aussage des Konfidenzintervalls nach (2.39) läßt sich für den normalverteilten Schätzer $\bar{X}$ präzisieren. Mit

$$A_1 = A_2 = z_{\alpha/2} \sqrt{\text{Var}(\bar{X})} \quad (3.26)$$

lautet das Konfidenzintervall dann

$$[\bar{X} \pm z_{\alpha/2} \sqrt{\text{Var}(\bar{X})}]. \quad (3.27)$$

Darin bezeichnet $z_{\alpha/2}$ die Realisation der standardnormalverteilten Zufallsvariablen Z^* , für die gilt

$$P(-z_{\alpha/2} \leq Z^* \leq z_{\alpha/2}) = \gamma = 1 - \alpha.$$

Bezüglich des Erwartungswertes des Schätzers bleibt die Aussage trotz Verzerrung erhalten: Das Zufallsintervall (3.27) überdeckt den Erwartungswert $E(\bar{X})$ mit der Wahrscheinlichkeit $\gamma = 1 - \alpha$, d. h. es gilt weiterhin

1) Vgl.: Bley Müller u. a., 83, S. 78.

$$P\left(\bar{X} - z_{\alpha/2} \sqrt{\text{Var}(\bar{X})} \leq E(\bar{X}) \leq \bar{X} + z_{\alpha/2} \sqrt{\text{Var}(\bar{X})}\right) = \gamma. \quad (3.28)$$

Bezogen auf den zu schätzenden wahren Parameter μ kann die Konfidenzaussage nicht mehr aufrecht erhalten werden, denn die Verteilung des Schätzers ist um $b_{\bar{X}}$ von μ weg verschoben. Die Wahrscheinlichkeit, daß der wahre Parameter in dem vorgegebenen Intervall um $\bar{X}$ liegt, sinkt mit zunehmender Verzerrung.

Um die Auswirkungen der Verzerrung $b_{\bar{X}}$ auf die Konfidenzaussage für μ abzuschätzen, ist zu berechnen, wie sich die Wahrscheinlichkeit $\gamma'(b_{\bar{X}})$ des mit (3.27) vorgegebenen Intervalls in Abhängigkeit von der Verzerrung ändert¹⁾:

$$P(\bar{X} - z_{\alpha/2} \sqrt{\text{Var}(\bar{X})} \leq \mu \leq \bar{X} + z_{\alpha/2} \sqrt{\text{Var}(\bar{X})}) = \gamma'(b_{\bar{X}}) = \gamma. \quad (3.29)$$

Die Wahrscheinlichkeit $\gamma'(b_{\bar{X}})$ zeigt auf, wie sich eine Vernachlässigung der systematischen Fehler auf die Schätzung auswirkt.

Mit $\mu = E(\bar{X}) - b_{\bar{X}}$ läßt sich die normalverteilte Zufallsvariable in (3.29) standardisieren, so daß mit (3.25) gilt:

$$\begin{aligned} \gamma'(b_{\bar{X}}) &= P\left(-z_{\alpha/2} \leq \frac{\bar{X} - \mu}{\sqrt{\text{Var}(\bar{X})}} \leq z_{\alpha/2}\right) \\ &= P\left(-z_{\alpha/2} \leq \frac{\bar{X} - E(\bar{X})}{\sqrt{\text{Var}(\bar{X})}} + b^* \leq z_{\alpha/2}\right) \\ &= P\left(-z_{\alpha/2} - b^* \leq Z^* \leq z_{\alpha/2} - b^*\right). \end{aligned} \quad (3.30)$$

Über die Verteilungsfunktion $F(0;1)$ der standardnormalverteilten Zufallsvariablen Z^* folgt dann

1) Vgl.: Cochran 72, S. 28 ff; Menges u. Skala, 72, S. 283 ff.

$$\gamma'(b_{\bar{x}}) = F(0;1)(z_{\frac{\alpha}{2}} - b^*) - F(0;1)(-z_{\frac{\alpha}{2}} - b^*). \quad (3.31)$$

Entscheidend für die Verringerung der Wahrscheinlichkeit des vorgegebenen Konfidenzintervalls ist die relative Verzerrung b^* , das Verhältnis zwischen Verzerrung und Standardabweichung des Schätzers.

Tabelle 3.2 gibt die Wahrscheinlichkeit γ' für verschiedene Konfidenzniveaus und ausgewählte Werte von b^* an. Wegen der Symmetrieeigenschaften der Normalverteilung genügt es, eine Tabelle für positive b^* anzugeben¹⁾.

b^*	$\gamma = 0,68$ $z_{\alpha/2} = 1,0$	$\gamma = 0,90$ ¹ $z_{\alpha/2} = 1,645$	$\gamma = 0,95$ $z_{\alpha/2} = 1,96$	$\gamma = 0,99$ $z_{\alpha/2} = 2,58$
0,00	0,68268	0,90003	0,95000	0,99011
0,05	0,68208	0,89961	0,94972	0,99002
0,10	0,68027	0,89833	0,94886	0,98974
0,15	0,67726	0,89622	0,94742	0,98928
0,20	0,67307	0,89325	0,94541	0,98862
0,25	0,66772	0,88945	0,94281	0,98776
0,30	0,66124	0,88480	0,93963	0,98670
0,40	0,64499	0,87301	0,93148	0,98393
0,50	0,62466	0,85792	0,92091	0,98020
0,60	0,60062	0,83960	0,90785	0,97541
0,75	0,55865	0,80629	0,88350	0,96594
1,00	0,47725	0,73645	0,82994	0,94278
1,50	0,30232	0,55681	0,67697	0,85992
2,00	0,15730	0,36115	0,48400	0,71905
2,50	0,06657	0,19624	0,29458	0,53188
3,00	0,02271	0,08768	0,14915	0,33722

Tabelle 3.2: Die tatsächliche Wahrscheinlichkeit γ' in Abhängigkeit von der relativen Verzerrung b^* ²⁾.

1) Die numerische Berechnung von $F(0;1)$ führt zu Rundungsfehlern in den angegebenen Werten; die ersten vier Stellen hinter dem Komma sind sicher.

2) Vgl.: Cochran, 72, S. 30; Menges u. Skala, 73, S. 285.

Die tatsächliche Wahrscheinlichkeit, mit der das Zufallsintervall (3.27) bei einem Konfidenzniveau von $\gamma = 0,95$, trotz einer relativen Verzerrung $b^* = 0,3$, den zu schätzende Parameter μ überdeckt, beträgt noch rund 0,94; bei einem $b^* = 1,0$ sinkt die tatsächliche Wahrscheinlichkeit bereits auf rund 0,83.

Die Auswirkungen der Verzerrung auf die Konfidenzaussage ist nicht linear; dies wird auch in Schaubild 3.2 deutlich.

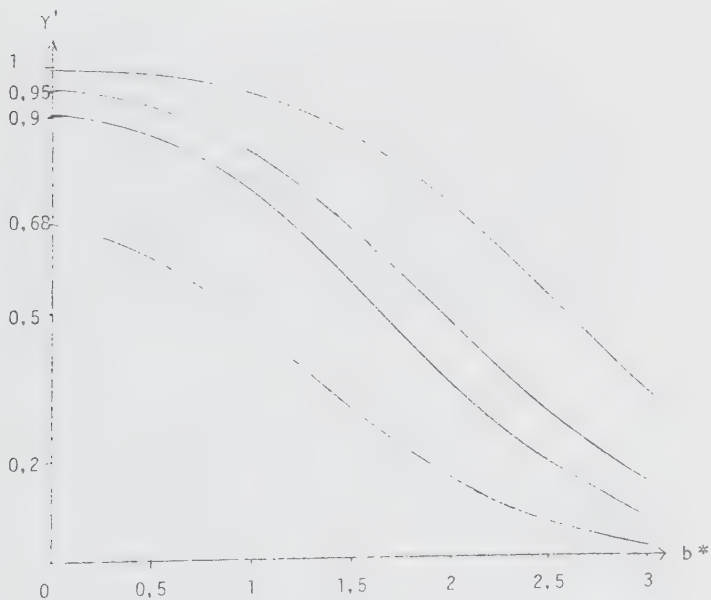


Schaubild 3.2: Die tatsächliche Wahrscheinlichkeit γ' in Abhängigkeit von der relativen Verzerrung b^* .

Die Berechnungen bestätigen im wesentlichen die Aussagen, die bereits durch die Betrachtung des Gesamtfehlers (vgl.: Abschnitt 3.2.3) getroffen wurden. Geringfügige Verzerrungen, z. B. $b^* < 0,25$, verringern die Wahrscheinlichkeit der Konfidenzaussage kaum, so daß eine Vernachlässigung gerechtfertigt erscheint.

3.2.4.2 Ein Versuch, die Auswirkungen der Verzerrung über den Stichprobenumfang zu verdeutlichen

Die Verzerrung des Schätzers $\bar{X}$ führt zu einer Verschiebung des Konfidenzintervalls. Dadurch sinkt, wie die Ausführungen des letzten Abschnitts zeigen, die tatsächliche Wahrscheinlichkeit, mit der das Zufallsintervall den zu schätzenden Parameter überdeckt, vom Konfidenzniveau γ auf γ' . Die Länge des Intervalls bleibt jedoch unbeeinflusst.

Eine gleichwertige Schätzaussage, d. h. ein Zufallsintervall gleicher Länge und einem Konfidenzniveau von γ' , wäre bei niedrigerem Stichprobenumfang möglich, wenn keine Verzerrung des Schätzers bestünde. Wie groß wäre die Ersparnis im Stichprobenumfang, die durch die bestehende Verzerrung des Schätzers verhindert ist?

Das mit dem Schätzer $\bar{X}$ formulierte Konfidenzintervall (3.27) lautet:

$$\bar{X} \pm z_{\frac{\gamma}{2}} \cdot \sqrt{\text{Var} \bar{X}}.$$

Die Einbeziehung des Stichprobenumfangs n in die Schätzungsaussage kann über $\text{Var}(\bar{X})$ erfolgen. Bei einfacher Zufallsauswahl mit Zurücklegen gilt speziell (vgl. (3.10))

$$\text{Var}(\bar{X}) = \frac{\sigma^2}{n},$$

so daß für das Konfidenzintervall folgt:

$$\bar{X} \pm z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}.$$

Soll zum Vergleich mit einer um b^* verzerrten Schätzung eine Konfidenzaussage mit einem Konfidenzniveau $\gamma = \gamma'(b^*)$ erfolgen, so ist aus der Verteilungsfunktion der standardisierten Normalverteilung ein $z_{\alpha'/2}$ mit $\alpha' = 1 - \gamma'$ zu bestimmen. Dieser Wert ist geringer als $z_{\alpha/2}$, weil $\gamma' < \gamma$ gilt. Um die daraus resultierende Reduzierung der Intervalllänge auszugleichen - es soll ja eine gleichwertige Aussage erfolgen - kann n auf n' verringert werden. σ ist mit der Gesamtheit vorgegeben. Daraus folgt:

$$\frac{z_{\alpha/2}}{\sqrt{n}} = \frac{z_{\alpha'/2}}{\sqrt{n'}}$$

$$\frac{n'}{n} = \left(\frac{z_{\alpha'/2}}{z_{\alpha/2}} \right)^2, \quad (3.32)$$

In Tabelle 3.3 sind die Werte $\frac{n'}{n}$ für verschiedene relative Verzerrungen b^* zusammengestellt. Dabei ist ein ursprüngliches Konfidenzniveau $\gamma = 0,95$ bzw. $\gamma = 0,9$ angenommen.

b*	$\gamma = 0,9$		$\gamma = 0,95$	
	$\frac{z_{\alpha'}}{2}$	$\frac{n'}{n}$	$\frac{z_{\alpha'}}{2}$	$\frac{n'}{n}$
0	1,64	1,000	1,96	1,000
0,25	1,59	0,934	1,90	0,940
0,5	1,47	0,799	1,76	0,806
1,0	1,12	0,464	1,37	0,490
1,5	0,75	0,208	0,99	0,255
2,0	0,47	0,082	0,65	0,110

Tabelle 3.3: Vergleich zwischen verzerrter und unverzerrter Schätzung über den Stichprobenumfang, bezogen auf die tatsächliche Wahrscheinlichkeit $\gamma'(b^*)$ (Vgl. Tabl. 3.2).

Gehen wir beispielsweise von einer relativen Verzerrung $b^* = 0,5$ aus, so beträgt die tatsächliche Wahrscheinlichkeit der Intervallschätzungsaussage bei einem Konfidenzniveau von 0,95 noch 0,92091. Eine vergleichbare Schätzungsaussage hätte bei unverzerrter Schätzung auch mit einem knapp 20 % niedrigeren Stichprobenumfang erreicht werden können (das Verhältnis $\frac{n'}{n}$ beträgt 0,806).

Bei höherer relativer Verzerrung wäre die Ersparnis des Stichprobenumfanges größer. Beträgt die relative Verzerrung $b^* = 1$, so hätte bei unverzerrter Schätzung bereits ein Stichprobenumfang von $\frac{n}{2}$ ausgereicht ($\frac{n'}{n} = 0,49$), um eine gleichwertige Schätzung zu ermöglichen.

Die Tabelle ist für die einfache Zufallsauswahl mit Zurücklegen ausgerechnet. Für das Ziehen ohne Zurücklegen verhindert die Endlichkeitskorrektur ein leichtes Umrechnen der Auswirkungen der Verzerrung in eine Aussage bezogen auf den notwendigen Stichprobenumfang. Ist der Auswahlssatz jedoch klein, was in der Praxis wegen der großen Gesamtheit gegeben sein dürfte, kann die Tabelle auf die Auswahl ohne Zurücklegen übertragen werden.

Der Vergleich einer verzerrten Schätzung mit einer unverzerrten über den notwendigen Stichprobenumfang zeigt nochmals deutlich, welche Bedeutung der Vermeidung bzw. Ermittlung des bias zukommt. Die Ergebnisse des Vergleichs lassen sich zur besseren Verteilung der für eine Stichprobenuntersuchung zur Verfügung stehenden finanziellen Mittel auf die Vermeidung bzw. Ermittlung der Verzerrung und den Stichprobenumfang nutzen.

Ist beispielsweise die Ersparnis durch die Reduzierung des Stichprobenumfangs größer als die Kosten der Ermittlung der Verzerrung, so sind die Mittel auf die Bestimmung des bias zu verwenden.

Zu beachten ist allerdings, daß sich der Vergleich auf die tatsächliche Wahrscheinlichkeit der verzerrten Intervallschätzung bezieht, nicht auf das vorgegebene Konfidenzniveau. Um dieses zu erreichen, ist ein Intervall um $\bar{X}$ zu bestimmen, das den wahren Parameter μ mit vorgegebener Wahrscheinlichkeit γ überdeckt, trotz der Verzerrung b^* . Dazu muß man ein $z_{\alpha^*/2}$ in Abhängigkeit von b^* berechnen für das gilt:

$$P\left(\bar{X} - z_{\frac{\alpha^*}{2}} \sqrt{\text{Var}(\bar{X})} \leq \mu \leq \bar{X} + z_{\frac{\alpha^*}{2}} \sqrt{\text{Var}(\bar{X})}\right) = \gamma^*(b^*) = \gamma. \quad (3.33)$$

Dieses $z_{\alpha^*/2}$ kann iterativ aus der Normalverteilungstabelle (Annahme: normalverteilter Schätzer $\bar{X}$) ermittelt werden. Die negativen Auswirkungen der Verzerrung b^* auf die Konfidenzaussage werden durch eine Verlängerung des Intervalls ($z_{\alpha^*/2}$ steigt mit b^*) ausgeglichen.

Welcher Stichprobenumfang n^* hätte ausgereicht, eine vergleichbare Konfidenzaussage (gleiche Intervalllänge, vorgegebenes γ) zu erhalten, wenn keine Verzerrung vorläge? Die Einbeziehung des Stichprobenumfangs erfolgt über $\text{Var}(\bar{X})$, wobei wiederum eine einfache Zu-

fallsauswahl mit Zurücklegen unterstellt ist¹⁾). Bei gegebener Varianz der Merkmalsausprägungen in der Gesamtheit ist die Intervalllänge nur durch $z_{\alpha^*/2}$ und $\sqrt{n}$ bestimmt, so daß analog zu (3.32) gilt:

$$\frac{z_{\alpha^*/2}}{\sqrt{n}} = \frac{z_{\alpha/2}}{\sqrt{n^*}}$$

$$\frac{n^*}{n} = \left(\frac{z_{\alpha/2}}{z_{\alpha^*/2}} \right)^2 \quad (3.34)$$

Mit der vorgegebenen Wahrscheinlichkeit γ ist $z_{\alpha/2}$ bestimmt - der Vergleich bezieht sich ja auf eine unverzerrte Schätzung.

In Tabelle 3.4 sind die Ergebnisse einer solchen Rechnung für ausgewählte b^* und die vorgegebenen Konfidenz-

	b^*	α^*	$z_{\alpha^*/2}$	$\frac{n^*}{n}$
$\gamma = 0,9:$	0,00	0,100	1,645	1,000
	0,25	0,090	1,696	0,941
	0,50	0,062	1,839	0,800
	1,00	0,022	2,285	0,518
	1,50	0,005	2,782	0,350
	2,00	0,001	3,282	0,251
$\gamma = 0,95:$	0,00	0,050	1,960	1,000
	0,25	0,043	2,020	0,941
	0,50	0,029	2,182	0,807
	1,00	0,008	2,647	0,548
	1,50	0,002	3,145	0,388
	2,00	0,000	3,645	0,289

Tabelle 3.4: Vergleich zwischen verzerrter und unverzerrter Schätzung, bezogen auf das Konfidenzniveau γ .

1) Die Endlichkeitskorrektur bei ohne Zurücklegen gezogenen Stichproben verhindert eine einfache Umrechnung. Die Werte können dann als Näherung gelten.

niveaus $\gamma = 0,9$ bzw. $\gamma = 0,95$ angegeben. Neben dem Verhältnis $\frac{n^*}{n}$ sind auch α^* und $z_{\alpha^*}/2$ mit aufgelistet.

Ist beispielsweise bei einer Erhebung mit einer relativen Verzerrung $b^* = 1$ zu rechnen und soll eine Intervallschätzung mit $\gamma = 0,95$ erfolgen, so hätte eine vergleichbare Schätzaussage bei unverzerrtem $\bar{X}$ mit einem Stichprobenumfang n^* erreicht werden können, der nur rund 55 % des gegebenen Stichprobenumfangs n beträgt. Bemerkenswert ist, daß bei geringerer vorgegebener Wahrscheinlichkeit (als Ausgleich der vermiedenen Verzerrung) eine stärkere Reduzierung des Stichprobenumfanges erfolgen kann. Das gleiche gilt für den Vergleich zwischen verzerrter und unverzerrter Schätzung,

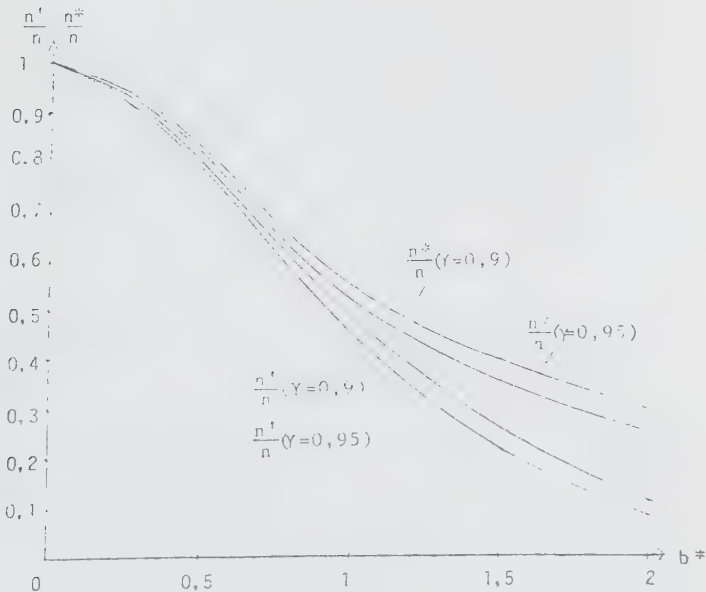


Schaubild 3.3: Vergleich zwischen verzerrter und unverzerrter Schätzung über den Stichprobenumfang.

bezogen auf die tatsächliche Wahrscheinlichkeit γ' . In Schaubild 3.3, in dem neben $\frac{n'}{n}$ auch $\frac{n^*}{n}$ in Abhängigkeit von b^* für die betrachteten Konfidenzniveaus dargestellt ist, kommt das darin zum Ausdruck, daß die Kurve für $\gamma = 0,90$ unterhalb der jeweiligen für $\gamma = 0,95$ verläuft.

Dies hängt damit zusammen, daß die Intervalle mit geringerem Konfidenzniveau stärker auf eine Verzerrung reagieren, weil ihre Intervallgrenzen in einem Bereich der Dichtefunktion liegen, der eine größere Steigung aufweist. Eine Parallelverschiebung der Intervallgrenzen durch die Verzerrung von $\bar{X}$ wirkt sich deshalb stärker auf die Wahrscheinlichkeit des Intervalls aus, die dann eine stärkere Reduzierung des Stichprobenumfanges als Ausgleich der Verzerrung ermöglicht.

Ein Vergleich der Ergebnisse $\frac{n^*}{n}$ mit $\frac{n'}{n}$ aus Tabelle 3.3 zeigt, daß die Reduzierung des Stichprobenumfanges zum Ausgleich der fehlenden Verzerrung bei vorgegebenem γ geringer ausfällt; die Kurven für $\frac{n'}{n}$ verlaufen unterhalb der entsprechenden für $\frac{n^*}{n}$ (vgl.: Schaubild 3.3). Die geringere tatsächliche Wahrscheinlichkeit γ' muß durch eine Steigerung des Stichprobenumfanges ausgeglichen werden.

3.2.5 Möglichkeiten einer Berücksichtigung der Verzerrung in der Schätzaussage

Deskriptive Kontrollmaßnahmen ermöglichen keine Korrektur der Schätzung, ihre Aufgabe liegt in der Beschreibung, Kommentierung der Ergebnisse. Die Resultate der deskriptiven Kontrollen müssen daher in der Veröffentlichung der Schätzergebnisse enthalten sein.

Wenn operative Kontrollmaßnahmen erfolgreich waren, die Verzerrung $b_{\bar{X}}$ ist dann bekannt, kann eine Korrektur der Mittelwertschätzung erfolgen¹⁾. Ein erwartungstreuer Schätzer für μ ist dann

$$\bar{X}_k = \bar{X} - b_{\bar{X}}; \quad (3.35)$$

$$\begin{aligned} E(\bar{X}_k) &= E(\bar{X}) - b_{\bar{X}} \\ &= \mu + b_{\bar{X}} - b_{\bar{X}} = \mu. \end{aligned} \quad (3.36)$$

Wird ein Konfidenzintervall mit diesem korrigierten Schätzer $\bar{X}_k$ formuliert, so bleibt die Aussage bezüglich μ erhalten, weil die Verschiebung der Verteilung von $\bar{X}$ wieder rückgängig gemacht wurde.

Die Ergebnisse der operativen Kontrollmaßnahmen lassen nicht immer definitive Aussagen über die Gesamtverzerrung zu, z. B. weil sie nur einzelne Fehlerquellen erfassen. Die Schätzung kann dann zwar um die ermittelte Verzerrung korrigiert werden, die möglicherweise bestehende zusätzliche Verzerrung bleibt dann unberücksichtigt.

Geben die Kontrollmaßnahmen Hinweise und Anhaltspunkte über die Restverzerrung, sind diese jedoch für eine Korrektur zu unsicher, so sollten die Informationen in der Schätzaussage berücksichtigt werden. Eine Korrektur der Schätzung ist dann zwar nicht möglich, allerdings kann man versuchen, die Sicherheit der Intervallaussage zu erhöhen. Es geht darum, die durch die Verzerrung der Schätzung entstehende Reduzierung der tatsächlichen Wahrscheinlichkeit der Konfidenzaussage auszugleichen.

Eine einfache Möglichkeit, die tatsächliche Wahrscheinlichkeit des Konfidenzintervalls zu steigern, besteht darin, ein sichereres Konfidenzniveau vorzugeben. Bei

1) Vgl.: Pokropp, 80, S. 212.

einer relativen Verzerrung von $b^* = 1$ ist beispielsweise bei $\gamma = 0,99$ noch mit einer tatsächlichen Wahrscheinlichkeit von 0,94 zu rechnen. Die Erhöhung der Sicherheit geht natürlich auf Kosten der Präzision; das Intervall wird länger.

Eine andere Möglichkeit besteht - zumindest theoretisch - darin, an Stelle der Standardabweichung des Schätzers den Gesamtfehler Gf bei der Bestimmung des Konfidenzintervalls zu benutzen¹⁾. Dadurch geht der Bias mit in die Schätzaussage ein.

Das modifizierte Konfidenzintervall lautet dann:

$$(\bar{X} \pm z_{\alpha/2} \text{ Gf}). \quad (3.37)$$

Der Gesamtfehler ist in Abschnitt 3.2.3 definiert als Quadratwurzel des Mean Square Error. Mit (3.23) und (3.24) läßt sich Gf umformen zu

$$\text{Gf} = \sqrt{\text{Var}(\bar{X})} \cdot c, \quad (3.38)$$

Die Grenzen des Intervalls (3.37) verändern sich mit b^* , denn im Faktor c ist die relative Verzerrung enthalten. Wegen dieser Intervallverlängerung ist die Wahrscheinlichkeit, mit der dieses modifizierte Intervall den wahren Parameter μ überdeckt, größer als $\gamma'(b_{\bar{X}}^*)$. Die Berechnung dieser Wahrscheinlichkeit kann in Analogie zu (3.31) erfolgen durch

$$\begin{aligned} \gamma'_{\text{Gf}} &= F(0;1) (z_{\alpha/2} c - b^*) \\ &\quad - F(0;1) (-z_{\alpha/2} c - b^*). \end{aligned} \quad (3.39)$$

1) Vgl.: Hansen u. a., I, 53, S. 56 ff.

b*	c	$\gamma = 0,683$	$\gamma = 0,9$	$\gamma = 0,95$	$\gamma = 0,99$
		$z_{\alpha/2}=1,0$	$z_{\alpha/2}=1,645$	$z_{\alpha/2}=1,96$	$z_{\alpha/2}=2,58$
0,00	1,000	0,68268	0,90003	0,95000	0,99011
0,05	1,001	0,68268	0,90003	0,95000	0,99011
0,10	1,005	0,68267	0,90003	0,95001	0,99012
0,15	1,011	0,68264	0,90003	0,95001	0,99012
0,20	1,020	0,68256	0,90002	0,95003	0,99015
0,25	1,031	0,68239	0,90001	0,95007	0,99019
0,30	1,044	0,68209	0,90000	0,95014	0,99027
0,40	1,077	0,68097	0,89999	0,95044	0,99057
0,50	1,118	0,67889	0,90008	0,95106	0,99109
0,60	1,166	0,67568	0,90042	0,95212	0,99184
0,75	1,250	0,66870	0,90176	0,95475	0,99330
1,00	1,414	0,65276	0,90721	0,96171	0,99595
1,50	1,803	0,61849	0,92863	0,97900	0,99918
2,00	2,236	0,59330	0,95338	0,99141	0,99990
2,50	2,693	0,57636	0,97318	0,99726	0,99998
3,00	3,162	0,56446	0,98619	0,99930	0,99999

Tabelle 3.5: Die tatsächliche Wahrscheinlichkeit γ'_{Gf} in Abhängigkeit von der relativen Verzerrung b^* ¹⁾.

Tabelle 3.5 gibt die tatsächliche Wahrscheinlichkeiten γ'_{Gf} des modifizierten Konfidenzintervalls für verschiedene Konfidenzniveaus γ in Abhängigkeit von der relativen Verzerrung b^* an. Die Intervallverlängerung geht aus dem in der Tabelle mit aufgeführten Faktor c hervor (vgl. auch Tabelle 3.1).

Bei einem Konfidenzniveau von $\gamma = 0,683$ ($z_{\alpha/2} = 1$; einfacher Varianzbereich bei normalverteiltem Schätzer) beträgt die Wahrscheinlichkeit bei $b^* = 0,5$ noch 0,679 gegenüber 0,625, wenn keine Veränderung der Intervall-

1) Vgl.: Hansen u. a., I. 53. S. 59

grenzen folgen würde. Das Intervall ist um rund 12 % ($c = 1,118$) länger.

Beträgt das Konfidenzniveau $\gamma = 0,90$ ($z_{\alpha/2} = 1,645$), bleibt γ'_{Gf} bis zu einer relativen Verzerrung von $b^* = 0,5$ in etwa konstant, d. h. die Intervallverlängerung gleicht die Verzerrung des Schätzers aus. Für größere b^* ist der Einfluß der Intervallverlängerung stärker, so daß die tatsächliche Wahrscheinlichkeit der Schätzaussage trotz der Verzerrung des Schätzers $\bar{X}$ sogar ansteigt. Bei modifizierten Konfidenzintervallen, die mit einem größeren $z_{\alpha/2}$ (größeres Konfidenzniveau) berechnet werden, steigt γ' schon für geringe relative Verzerrungen an.

Mit dem modifizierten Konfidenzintervall (3.37) läßt sich keine bestimmte vorgegebene Wahrscheinlichkeit γ einhalten, mit der das verzerrte Intervall den wahren Parameter μ überdeckt. Diese Forderung erfüllt wegen (3.33) das Intervall

$$[\bar{X} \pm z_{\alpha^*/2} \sqrt{\text{Var}(\bar{X})}].$$

Tabelle 3.4 kann man daher auch dazu nutzen, die Verzerrung in der Schätzaussage zu berücksichtigen.

Soll beispielsweise mit dem Schätzer $\bar{X}$ eine Intervallaussage mit $\gamma = 0,95$ bezüglich μ erfolgen, die Kontrollmaßnahmen geben aber Hinweise, daß eine relative Verzerrung $b^* = 0,5$ vorliegt, so kann das Konfidenzintervall mit einem $z_{\alpha^*/2} = 2,182$ formuliert werden. Dieses Zufallsintervall überdeckt den wahren Parameter μ mit der Wahrscheinlichkeit $\gamma = 0,95$.

Die Versuche, eine Verzerrung des Schätzers im Konfidenzintervall zu berücksichtigen, führen jeweils zu einem längeren Zufallsintervall. Nur dann, wenn die Verzerrung vermieden bzw. wenn sie ermittelt wurde, kann diese Verringerung der Schätzgenauigkeit umgangen werden.

3.3 Die Verzerrung beim Schichtenauswahlverfahren

3.3.1 Das Schichtenauswahlverfahren

Das Schichtenauswahlverfahren bietet gegenüber der einfachen Zufallsauswahl einige Vorteile¹⁾. Durch die Schichteneinteilung ist es möglich, aus Teilgesamtheiten Informationen zu erhalten. Zusätzlich kann die Organisation der Erhebung besser auf die Besonderheiten in den einzelnen Teilgesamtheiten eingehen.

Neben diesen praktischen Vorteilen, die im Einzelfall ausschlaggebend für die Wahl des Stichprobenverfahrens sein können, weist die geschichtete Auswahl eine genauere Schätzung auf. Der Zufallsfehler der Schätzung, gemessen durch die Varianz des Schätzers, ist im Vergleich zur einfachen Zufallsauswahl geringer. Dieser Aspekt wird in Abschnitt 3.4 detaillierter behandelt. Vorher ist das Modell der geschichteten Auswahl²⁾ zu beschreiben und anschließend sind systematische Fehler in diesem Modell zu berücksichtigen.

3.3.1.1 Grundlegende Notationen

Gegeben sei, wie bisher, die Grundgesamtheit G mit N Elementen. Diese ist in m Teilgesamtheiten G_h , $h = 1, 2, \dots, m$, Schichten genannt, eingeteilt mit den Umfängen N_h , $h = 1, 2, \dots, m$,

$$G_h = \{u_{hj} \mid j = 1, 2, \dots, N_h\}. \quad (3.40)$$

1) Vgl.: Cochran, 72, S. 111 f; Menges u. Skala, 73, S. 102.

2) Zum Modell der Schichtenauswahl vgl.: Cochran, 72, S. 111 ff; Menges u. Skala, 73, S. 101 ff; Pokropp, 30, S. 48 ff.

Auf das Problem der Schichteneinteilung: Anzahl der Schichten, Bestimmung der Schichtgrenzen, soll hier nicht weiter eingegangen werden¹⁾. Die Einteilung muß allerdings elementfremd erfolgen, so daß gilt

$$G_h \cap G_k = \emptyset \text{ für } h \neq k$$

und damit

$$\sum_h N_h = n.$$

Das Merkmal X ordnet jeder Untersuchungseinheit u_{hj} eine Merkmalsausprägung

$$x_{hj} = X(u_{hj}), \quad x_{hj} \in \mathbb{R}, \quad (3.41)$$

zu. In jeder Schicht existiert dann ein Mittelwert μ_h der Merkmalsausprägungen, also

$$\mu_h = \frac{1}{N_h} \sum_j x_{hj}. \quad (3.42)$$

Das arithmetische Mittel μ der Merkmalsausprägungen in der Gesamtheit G ergibt sich mit dieser Schreibweise

$$\mu = \frac{1}{N} \sum_h \sum_j x_{hj} \quad (3.43)$$

oder mit (3.42)

$$\mu = \sum_h \frac{N_h}{N} \mu_h. \quad (3.44)$$

In (3.44) ist μ als gewogenes arithmetisches Mittel der Schichtenmittelwerte ausgedrückt, wobei die rela-

1) Zum Problem der Schichteneinteilung vgl.: Statistisches Bundesamt, 60, S. 74 ff; Dalenius u. Hodges, 57, S. 198 ff.

tiven Schichtenumfänge $\frac{N_h}{N} = W_h$ als Gewichtungsfaktoren fungieren. Für die W_h gilt

$$0 < W_h < 1, \quad \sum_h W_h = 1.$$

Analog zur einfachen Zufallsauswahl existiert in jeder Schicht eine Varianz der Merkmalsausprägungen

$$\sigma_h^2 = \frac{1}{N_h} \sum_j (x_{hj} - \mu_h)^2 \quad (3.45)$$

bzw. in der korrigierten Form

$$s_h^2 = \frac{N_h}{N_h - 1} \sigma_h^2 = \frac{1}{N_h - 1} \sum_j (x_{hj} - \mu_h)^2. \quad (3.46)$$

Die Varianz des Merkmals X in der Gesamtheit

$$\sigma^2 = \frac{1}{N} \sum_h \sum_j (x_{hj} - \mu)^2 = \frac{N-1}{N} s^2$$

kann aufgespalten werden¹⁾ in

$$\sigma^2 = \sum_h W_h \sigma_h^2 + \sum_h W_h (\mu_h - \mu)^2. \quad (3.47)$$

Der erste Term stellt dabei die mit den relativen Schichtenumfängen gewichtete durchschnittliche Varianz innerhalb der Schichten dar. Der zweite Term

$$\sigma_z^2 = \sum_h W_h (\mu_h - \mu)^2$$

wird als Varianz zwischen den Schichten bezeichnet. Ähnliches gilt auch für die korrigierte Varianz²⁾

$$s^2 = \frac{1}{N-1} \sum_h (N_h - 1) s_h^2 + \frac{1}{N-1} \sum_h N_h (\mu_h - \mu)^2. \quad (3.48)$$

1) Vgl.: z. B. Menges u. Skala, 73, S. 103.

2) Vgl.: Cochran, 72, S. 123 f.

3.3.1.2 Die Mittelwertschätzung bei geschichteter Auswahl

Der Mittelwert μ der Gesamtheit soll mit Hilfe einer Stichprobe geschätzt werden. Dazu wird aus jeder Schicht unabhängig von den anderen Schichten eine Zufallsstichprobe vom Umfang n_h gezogen. Der Gesamtstichprobenumfang n ist dann

$$n = \sum_h n_h. \quad (3.49)$$

Wie bisher unterstellen wir, daß die befragten Individuen mit ihren wahren Merkmalsausprägungen antworten - Antwortfehler bleiben in diesem Kapitel weiterhin unberücksichtigt.

Jede Schicht ist eine Teilgesamtheit, aus der eine einfache Zufallsstichprobe gezogen wird. Es gelten daher die Ausführungen für die einfache Zufallsauswahl in jeder Schicht. Zur Identifizierung der einzelnen Schichten ist jeweils der Index h hinzuzufügen. Aus Schicht h wird analog zu (2.10) die Stichprobe

$$\underline{X}_h = (X_{hi}/i = 1, 2, \dots, n_h) \quad (3.50)$$

gezogen. Die X_{hi} sind Zufallsvariablen mit

$$G_{Xh} = (x_{hj}/j = 1, 2, \dots, N_h)$$

als Folge der Realisationen, wobei die einzelnen Realisationen mit gleicher Wahrscheinlichkeit

$$P(X_{hi} = x_{hj}) = \frac{1}{N_h}, \quad j = 1, 2, \dots, N_h,$$

angenommen werden. Das arithmetische Mittel

$$\bar{X}_h = \frac{1}{n_h} \sum_i X_{hi} \quad (3.51)$$

ist wegen (3.9) ein erwartungstreuer Schätzer für μ_h , sofern in Schicht h keine systematischen Fehler auftreten.

Die Varianz des Mittelwertschätzers aus Schicht h beträgt nach (3.14)

$$\text{Var}(\bar{X}_h) = \frac{1}{n_h} \sigma_h^2 (1 - f_h^* a_h) \quad (3.52)$$

mit $f_h^* = \frac{n_h - 1}{N_h - 1}$.

In a_h kommt die bei der Varianz des Schätzers notwendige Unterscheidung der Auswahlmodi zum Tragen: $a_h = 0$ falls die Stichprobe in Schicht h mit Zurücklegen gezogen wurde; $a_h = 1$ falls die Stichprobe in Schicht h ohne Zurücklegen gezogen wurde.

Mit

$$\widehat{\text{Var}}(\bar{X}_h) = \frac{1}{n_h} S_h^2 (1 - f_h a_h), \quad f_h = \frac{n_h}{N_h}, \quad (3.53)$$

ist in Analogie zu (3.16) ein erwartungstreuer Schätzer für die Varianz des Mittelwertschätzers aus Schicht h gegeben, wobei S_h^2 die korrigierte Stichprobenvarianz

$$S_h^2 = \frac{1}{n_h - 1} \sum_h (X_{hi} - \bar{X}_h)^2$$

und f_h der Auswahlanteil der Stichprobe in Schicht h bedeutet.

Wegen der Erwartungstreue von $\bar{X}_h$ bezüglich μ_h läßt sich aus (3.44) ein erwartungstreuer Schätzer für μ ableiten. Es gilt

$$\bar{X}_g = \sum_h w_h \bar{X}_h; \quad (3.54)$$

$$\begin{aligned} E(\bar{X}_g) &= \sum_h w_h E(\bar{X}_h) \\ &= \sum_h w_h \mu_h = \mu. \end{aligned}$$

Die Schätzer $\bar{X}_h$ der einzelnen Schichten werden durch die relative Schichtengröße W_h auf die Gesamtheit hochgerechnet. Weil die Teilstichproben in den einzelnen Schichten getrennt voneinander gezogen werden, sind die Schätzer $\bar{X}_h$ unabhängige Zufallsvariablen und es gilt

$$\begin{aligned}\text{Var}(\bar{X}_g) &= \text{Var}(\sum_h W_h \bar{X}_h) \\ &= \sum_h W_h^2 \text{Var}(\bar{X}_h).\end{aligned}$$

Unter Beachtung von (3.52) folgt weiter

$$\text{Var}(\bar{X}_g) = \sum_h W_h^2 \frac{1}{n_h} \sigma_h^2 (1 - f_h a_h). \quad (3.55)$$

Da jede Schicht zur Schätzung des Parameters μ einen Beitrag liefert, hängt die Varianz des Schätzers $\bar{X}_g$ nur von den Varianzen innerhalb der Schichten, nicht aber von der Varianz zwischen den Schichten ab.

Die Schätzung von $\text{Var}(\bar{X}_g)$ erfolgt mit den korrigierten Stichprobenvarianzen S_h^2 aus den Schichten. Wegen (3.53) ist der Schätzer

$$\widehat{\text{Var}}(\bar{X}_g) = \sum_h \frac{1}{n_h} W_h^2 S_h^2 (1 - f_h a_h) \quad (3.56)$$

erwartungstreu bezüglich $\text{Var}(\bar{X}_g)$. Gleichung (3.55) zeigt, daß die Varianz des Schätzers im wesentlichen von der Schichteneinteilung (über σ_h^2 und W_h) und von der Verteilung der Stichprobe auf die Schichten (n_h) abhängt. In der Literatur¹⁾ werden mehrere Aufteilungsmöglichkeiten des Gesamtstichprobenumfangs auf die Schichten diskutiert:

(1) proportionale Aufteilung; Die Aufteilung erfolgt entsprechend den relativen Schichtenumfängen. Für

1) Vgl.: z. B. Menges u. Skala, 73, S. 105 ff; Cochran, 72, S. 116 f u. S. 120 ff; Hansen u. a., II, 53, S. 132 ff.

n_h ergibt sich:

$$n_{h, \text{prop}} = n W_h, \quad h = 1, 2, \dots, m.$$

(2) optimale Aufteilung nach Tschuprow und Neyman:
Bei vorgegebenen Schichtgrenzen (d. h. W_h und σ_h^2 sind gegeben) werden die n_h bestimmt, für die $\text{Var}(\bar{X}_g)$ minimal wird. Für den Auswahlmodus 'Ziehen mit Zurücklegen' ergibt sich

$$n_{h, \text{opt}} = n \frac{N_h \sigma_h}{\sum_h N_h \sigma_h}.$$

(3) optimale Aufteilung unter Berücksichtigung der Erhebungskosten: Unter Beachtung der gegebenen Schichteneinteilung werden die n_h bestimmt, die für eine bestimmte Genauigkeitsanforderung die Erhebungskosten (lineare Kostenfunktion) minimieren bzw. bei vorgegebenen Erhebungskosten die Varianz des Schätzers minimieren.

Durch Einsetzen der jeweiligen n_h in (3.55) lassen sich spezielle Formeln für die Varianz des geschichteten Schätzers ableiten. In dieser Arbeit soll darauf verzichtet werden.

Die Varianz $\text{Var}(\bar{X}_g)$ ist ein Maß für den Zufallsfehler des geschichteten Schätzers $\bar{X}_g$, und wegen (3.56) läßt sich diese aus der Stichprobe schätzen. Neben der Abschätzbarkeit sind auch die anderen Eigenschaften der Zufallsfehler, analog zur einfachen Zufallsauswahl (Vgl.: Abschnitt 3.2.1.3), erfüllt. Auf eine eingehendere Betrachtung kann daher verzichtet werden.

3.3.2 Die Beschreibung systematischer Fehler im Schichtenauswahlverfahren

Im Modell des Schichtenauswahlverfahrens haben systematische Fehler ähnliche Ursachen und Auswirkungen wie im Modell der einfachen Zufallsauswahl. Es empfiehlt sich daher, an die Ausführungen in Abschnitt 3.2.2 anzuknüpfen. Zwei Verzerrungsursachen wurden dort näher betrachtet: fehlerhafte Auswahlgrundlagen und systematische Fehler im Auswahlprozeß.

Fehler in der Auswahlgrundlage und im Auswahlprozeß der Stichprobe schlagen sich auf die Ergebnisse der Teilschätzungen aus den einzelnen Schichten nieder. In den Teilstichproben $\underline{x}_h$, die unabhängig voneinander nach dem Modell der einfachen Zufallsauswahl gezogen werden, ist mit unterschiedlichen Verzerrungen b_h zu rechnen. Es ist durchaus denkbar, daß nur in einzelnen Schichten die systematischen Fehler auftreten, z. B. weil nur Elemente aus bestimmten Schichten in der Auswahlgrundlage fehlen oder weil Fehler im Auswahlprozeß nur in einzelnen Schichten auftreten. Gerade weil das Schichtenauswahlverfahren eine Anpassung der Untersuchung an erhebungstechnische Besonderheiten der einzelnen Schichten ermöglicht, sind unterschiedliche Schichtenverzerrungen zu erwarten.

Da die Planung einer Erhebung die Gesamtheit der Schichten umfaßt, ist es durchaus möglich, daß einzelne Fehlerursachen in allen Schichten wirken. Die Schätzer für die Schichtenmittelwerte können daher, trotz der stochastischen Unabhängigkeit der $\bar{x}_h$, alle in die gleiche Richtung verzerrt sein.

Im folgenden seien Verzerrungen b_h der Schätzer $\bar{x}_h$ für den Schichtenmittelwert μ_h unterstellt. Der Erwartungswert von $\bar{x}_h$ beträgt dann:

$$E(\bar{X}_h) = \mu_h + b_h, \quad h = 1, 2, \dots, m. \quad (3.57)$$

Die Verzerrung des Schätzers $\bar{X}_g$ läßt sich unter diesen Voraussetzungen präzisieren:

$$\begin{aligned} E(\bar{X}_g) &= E\left(\sum_h W_h \bar{X}_h\right) \\ &= \sum_h W_h E(\bar{X}_h) \\ &= \sum_h W_h (\mu_h + b_h) \\ &= \sum_h W_h \mu_h + \sum_h W_h b_h \\ &= \mu + \sum_h W_h b_h. \end{aligned}$$

Eine Verzerrung innerhalb der einzelnen Schichten führt zu einer Gesamtverzerrung $b_{\bar{X}_g}$ als gewogenes arithmetisches Mittel der Schichtenverzerrungen b_h ,

$$b_{\bar{X}_g} := \sum_h W_h b_h. \quad (3.58)$$

Die Schätzung $\bar{X}_g$ ist nicht mehr erwartungstreu bezüglich μ , denn die Verteilung des Schätzers ist um $b_{\bar{X}_g}$ verschoben.

Zu beachten ist, daß bei verzerrter Auswahlgrundlage die tatsächlichen relativen Schichtenumfänge W_h unbekannt sind. Bekannt sind lediglich die Schichtenumfänge (relativ und absolut) der der vorhandenen Auswahlgrundlage zugrundeliegenden verzerrten Gesamtheit.

Die Varianz des Schätzers bzw. die Varianz des Zufallsfehlers ist mit Gleichung (3.55) angegeben, wobei zu beachten ist, daß systematische Fehler auch Einfluß auf $\text{Var}(\bar{X}_g)$ haben können. Z. B. hängt die Varianz des Schätzers aus einer verzerrten Auswahlgrundlage von der Varianz des Merkmals in der zugrundeliegenden verzerrten Gesamtheit ab, nicht aber von der Varianz in der tatsächlichen, echten Gesamtheit.

Entsprechend den Ausführungen in Abschnitt 3.2.3 läßt sich für den geschichteten Schätzer $\bar{X}_g$ der Gesamtfehler bzw. der Mean Square Error definieren und in einem Fehlerdreieck darstellen. Auf eine detaillierte Betrachtung kann jedoch verzichtet werden.

3.3.3 Auswirkungen der Verzerrung auf die Schätz- aussage beim Schichtenauswahlverfahren

3.3.3.1 Das Konfidenzintervall bei verzerrtem ge- schichtetem Schätzer

Der Ansatz, die Ergebnisse einer Erhebung in einem Konfidenzintervall zu formulieren, ist unabhängig von gewählten Stichprobenverfahren, sofern die Auswahl auf dem Zufallsprinzip beruht. Auch für den geschichteten Schätzer läßt sich daher ein Konfidenzintervall formulieren.

Für genügend große Stichprobenumfänge in den Schichten kann man, wie bei der einfachen Zufallsauswahl, für $\bar{X}_g$ Normalverteilung voraussetzen, wobei die Parameter der Verteilung mit $E(\bar{X}_g)$ und $\text{Var}(\bar{X}_g)$ gegeben sind. Das Konfidenzintervall lautet

$$[\bar{X}_g \pm z_{\alpha/2} \sqrt{\text{Var}(\bar{X}_g)}] \quad (3.59)$$

und es überdeckt mit der Wahrscheinlichkeit $\gamma = 1 - \alpha$ den zu schätzenden Parameter μ , sofern keine systematischen Fehler in der Schätzung vorliegen.

Die Auswirkungen der Verzerrung auf die Aussage des Konfidenzintervalls bei geschichteter Schätzung sind denen der einfachen Zufallsauswahl ähnlich, denn die Bildung des Intervalls erfolgt analog. Die Wahrscheinlichkeit

$$P\left\{\bar{X}_g - z_{\alpha/2} \sqrt{\text{Var}(\bar{X}_g)} \leq \mu \leq \bar{X}_g + z_{\alpha/2} \sqrt{\text{Var}(\bar{X}_g)}\right\} = \gamma'_g$$

sinkt mit zunehmender Verzerrung $b_{\bar{X}_g}$. Sie kann (vgl. (3.31)) über

$$\begin{aligned} \gamma'_g &= F(0,1) \left[z_{\alpha/2} - \frac{b_{\bar{X}_g}}{\sqrt{\text{Var}(\bar{X}_g)}} \right] \\ &\quad - F(0,1) \left[-z_{\alpha/2} - \frac{b_{\bar{X}_g}}{\sqrt{\text{Var}(\bar{X}_g)}} \right] \end{aligned} \quad (2.60)$$

berechnet werden. Entsprechend der einfachen Zufallsauswahl ist auch bei geschichteter Auswahl die relative Verzerrung

$$b_g^* = \frac{b_{\bar{X}_g}}{\sqrt{\text{Var}(\bar{X}_g)}}, \quad (3.61)$$

das Verhältnis zwischen Verzerrung und Standardabweichung des Schätzers $\bar{X}_g$, bestimmend für γ'_g . Die Tabelle 3.2 aus Abschnitt 3.2.4.1 läßt sich daher auch für die geschichtete Auswahl anwenden, wenn $b_g^* = b^*$ gilt.

Die Faustregel, für welche relative Verzerrung der Einfluß auf γ'_g so gering ist, daß er vernachlässigt werden kann, behält auch für die geschichtete Auswahl ihre Gültigkeit. Allerdings weist die Varianz des Schätzers beim Schichtenauswahlverfahren oft einen geringeren Wert als bei der einfachen Zufallsauswahl mit gleichem Stichprobenumfang n auf, so daß sich gleiche absolute Verzerrungen beim Schichtenauswahlverfahren stärker auswirken. Im Abschnitt 3.4 werden die Auswirkungen einer konstanten Verzerrung auf beide Verfahren detaillierter verglichen.

3.3.3.2 Ein Vergleich zwischen verzerrter und unverzerrter geschichteter Schätzung über den Stichprobenumfang

Im Zusammenhang mit den Auswirkungen einer Verzerrung auf die Schätzaussage bei einfacher Zufallsauswahl wird untersucht, welcher Stichprobenumfang ausgereicht hätte, eine vergleichbare Konfidenzaussage (gleiche Intervalllänge) zu erhalten, wenn die Verzerrung der Schätzung nicht bestünde. Der Vergleich wurde auf die tatsächliche Wahrscheinlichkeit des verzerrten Konfidenzintervalls und auf ein vorgegebenes Konfidenzniveau γ bezogen.

Liegt eine geschichtete Stichprobenerhebung vor, so erfordert diese Umrechnung eine Annahme über die Aufteilung des Stichprobenumfangs auf die einzelnen Schichten. Nur dann ist die Varianz des geschichteten Schätzers direkt abhängig vom Gesamtstichprobenumfang. Für den einfacheren Fall mit Zurücklegen gezogener Stichproben gilt für proportionale bzw. optimale Aufteilung¹⁾

$$\text{Var}(\bar{x}_g)_{\text{prop}} = \frac{1}{n} \sum_h W_h \sigma_h^2$$

bzw.

$$\text{Var}(\bar{x}_g)_{\text{opt}} = \frac{1}{n} \left(\sum_h W_h \sigma_h \right)^2.$$

In das Konfidenzintervall (3.59) geht daher auch bei geschichteter Auswahl der Faktor $1/\sqrt{n}$ ein, so daß analog zur einfachen Zufallsauswahl über die Beziehung (3.32) bzw. (3.34) eine Umrechnung der Verzerrung in den Stichprobenumfang erfolgen kann. Die Tabellen 3.3 und 3.4 sind daher auch für die geschichtete Auswahl mit proportionaler bzw. optimaler Aufteilung anwendbar.

1) Vgl.: z. B. Cochran, 72, S. 115 f u. S. 22.

3.4 Vergleich zwischen einfacher Zufallsauswahl und Schichtenauswahl

3.4.1 Der Schichteneffekt

Ziel dieses Abschnitts ist es, Unterschiede in den Auswirkungen einer Verzerrung auf die Untersuchungsergebnisse einer Erhebung mit geschichteter Auswahl und einfacher Zufallsauswahl aufzuzeigen.

Zunächst ist zu untersuchen, welches Verfahren eine bessere Schätzung liefert. Dazu bietet sich ein Vergleich der Varianzen der jeweiligen Schätzer an¹⁾.

Ein solcher Vergleich erfolgt über den Schichteneffekt; er stellt die Zufallsfehler der Schätzungen gegenüber.

Um den Vergleich von anderen Einflüssen unabhängig zu gestalten, sei anfangs unterstellt, daß keine systematischen Fehler in der Schätzung vorliegen.

Anschließend muß untersucht werden, wie systematische Fehler in den Schätzungen den Vergleich beeinflussen. Es ist die Frage zu beantworten, welches Verfahren empfindlicher auf eine Verzerrung reagiert. Bei diesem Vergleich spielt der Schichteneffekt eine bedeutende Rolle.

Natürlich sind auch praktische Aspekte bedeutsam, wenn Entscheidungen über das zu wählende Stichprobenverfahren anstehen. Fragen über Schichtenumfang, Schichtengrenzen, Anzahl der Schichten, Organisation der Erhebung, geforderte Informationen über Teilgesamtheiten beeinflussen in starkem Maß die Wahl des Erhebungsverfahrens. In einem theoretischen Vergleich der Schätzungen bleiben diese Kriterien jedoch unberücksichtigt.

1) Vgl.: Menges u. Skala, 73, S. 104 f; Cochran, 72, S. 123 ff.

Der Schichteneffekt D ist definiert als Differenz zwischen der Varianz des Schätzers bei einfacher Zufallsauswahl und der Varianz des geschichteten Schätzers. Für die Mittelwertschätzer $\bar{X}$ bzw. $\bar{X}_g$ gilt deshalb speziell¹⁾:

$$D = \text{Var}(\bar{X}) - \text{Var}(\bar{X}_g). \quad (3.62)$$

Ein positiver Schichteneffekt ($D > 0$) bedeutet daher, daß die geschichtete Auswahl eine präzisere Schätzung liefert als die einfache Zufallsauswahl.

Durch Einsetzen der Formeln für die Varianzen der jeweiligen Schätzer in (3.62) läßt sich der Schichteneffekt umformen. Wir beschränken uns auf den Auswahlvorgang 'Ziehen mit Zurücklegen'. Das vereinfacht die Betrachtungen erheblich, ohne daß die Ergebnisse für den Fall 'Ziehen ohne Zurücklegen' unbrauchbar wären. Mit (3.14) und (3.55) folgt für $a = a_h = 0$:

$$D = \frac{1}{n} \sigma^2 - \sum_h w_h^2 \frac{1}{n_h} \sigma_h^2. \quad (3.63)$$

Der besseren Vergleichbarkeit wegen und um sicherzustellen, daß die Unterschiede in den Varianzen auf das Verfahren und nicht auf die unterschiedlichen Stichprobenumfänge zurückzuführen sind, sei unterstellt, daß der Stichprobenumfang für beide Verfahren gleich ist, d. h. es gilt

$$n = \sum_h n_h.$$

Beide Verfahren beziehen sich auf die gleiche Gesamtheit. Die Verbindung zwischen den Varianzen innerhalb der Schichten und der Varianz der Gesamtheit stellt die Varianzaufspaltung nach (3.47) her:

1) Vgl.: Menges u. Skala, 73, S. 104 f.

$$\sigma^2 = \sum_h W_h \sigma_h^2 + \sigma_z^2.$$

Dies eingesetzt in (3.63) ergibt¹⁾

$$\begin{aligned} D &= \frac{1}{n} \sum_h W_h \sigma_h^2 + \frac{1}{n} \sigma_z^2 - \sum_h W_h^2 \frac{1}{n_h} \sigma_h^2 \\ &= \sum_h W_h \sigma_h^2 \left(\frac{1}{n} - \frac{1}{n_h} W_h \right) + \frac{1}{n} \sigma_z^2 \end{aligned}$$

und wegen $W_h = \frac{N_h}{N}$

$$D = \frac{1}{N^2} \sum_h N_h \sigma_h^2 \left(\frac{N}{n} - \frac{N_h}{n_h} \right) + \frac{1}{n} \sigma_z^2. \quad (3.64)$$

Bei proportionaler Aufteilung der Stichprobe auf die Schichten gilt

$$\frac{N}{n} = \frac{N_h}{n_h}, \quad h = 1, 2, \dots, m,$$

und daher folgt

$$D_{\text{prop}} = \frac{1}{n} \sigma_z^2 \geq 0. \quad (3.65)$$

Nur dann, wenn alle Schichtenmittelwerte μ_h gleich sind, ist die Varianz zwischen den Schichten $\sigma_z^2 = 0$ und dann sind Schichtenauswahlverfahren bei proportionaler Aufteilung und einfache Zufallsauswahl gleichwertig. In der Regel ist daher ein Schichtengewinn vorhanden.

Immer dann, wenn es gelingt, die Schichten so abzugrenzen, daß sich stark unterschiedliche μ_h ergeben, ist der Schichtengewinn besonders groß. Die Schichtung der Gesamtheit soll so erfolgen, daß die Schichten in sich möglichst homogen, untereinander jedoch möglichst inhomogen sind²⁾.

1) Vgl.: Menges u. Skala, 73, S. 104 f.

2) Vgl.: Menges u. Skala, 73, S. 105.

Für die optimale Aufteilung der Stichprobe auf die Schichten nach Tschuprow und Neyman, d. h. wenn $\text{Var}(\bar{X}_g)$ in Abhängigkeit von n_h minimal ist, ist der Schichteneffekt mindestens so groß wie D_{prop} . Es gilt¹⁾

$$D_{\text{opt}} = D_{\text{prop}} + \frac{1}{n} \sum_h W_h (\sigma_h - \sum_h W_h \sigma_h)^2. \quad (3.66)$$

Die optimale Aufteilung ist dann effektiv, wenn die Varianzen innerhalb der Schichten stark voneinander abweichen, denn den zweiten Summanden in (3.66) kann man als Varianz der Standardabweichungen σ_h interpretieren.

Mit dem Schichteneffekt (3.62) läßt sich die Varianz des geschichteten Schätzers auch schreiben als

$$\text{Var}(\bar{X}_g) = \text{Var}(\bar{X}) \left(1 - \frac{D}{\text{Var}(\bar{X})} \right). \quad (3.67)$$

Der zweite Faktor

$$d^2 = \left(1 - \frac{D}{\text{Var}(\bar{X})} \right) \quad (3.68)$$

bestimmt, um wieviel die Varianz der Schätzung durch die Schichtung sinkt. Entscheidend ist der relative Schichteneffekt

$$D^* = \frac{D}{\text{Var}(\bar{X})}. \quad (3.69)$$

Welche Werte D^* in der Praxis annehmen kann, hängt von der Qualität der Schichteneinteilung ab. Theoretisch sind die Grenzen für D^* durch

$$0 \leq D^* < 1$$

gegeben, sofern die Stichprobe proportional oder besser auf die Schichten verteilt ist.

1) Vgl.: Cochran, 72, S. 123 f.

3.4.2 Ein Vergleich zwischen einfacher Zufallsauswahl und Schichtenauswahl bei verzerrten Schätzungen

3.4.2.1 Die Verzerrung bei einfacher und geschichteter Zufallsauswahl

Ein Vergleich der geschichteten Auswahl mit der einfachen Zufallsauswahl unter Berücksichtigung der systematischen Fehler muß die Verzerrung der jeweiligen Schätzer erfassen. Es genügt nicht mehr, nur die Zufallsfehler über die Varianzen der Schätzer gegenüberzustellen.

Bevor man vergleichen kann, wie empfindlich die Schätzaussagen der beiden Verfahren auf einen bias reagieren, müssen die Verzerrungen der beiden Schätzer näher betrachtet werden. Da wir weiterhin korrekte individuelle Antworten voraussetzen, ist eine Beschränkung auf Fehler in der Auswahlgrundlage und im Auswahlprozeß als wichtige Verzerrungsursachen möglich¹⁾.

Die einfache Zufallsauswahl setzt eine Auswahlgrundlage voraus, die der Zielgesamtheit entspricht. Das Schichtenauswahlverfahren erfordert zusätzlich Informationen, die es erlauben, die Elemente der Gesamtheit den verschiedenen Schichten zuzuordnen, d. h. die Auswahlgrundlage in Schichten einzuteilen. Sind diese Informationen fehlerhaft, so wirkt sich das nicht als Verzerrung des Schätzers $\bar{X}_g$ aus, lediglich die Schichtung der Gesamtheit ist beeinträchtigt²⁾. Diese fehlerhafte Schichteneinteilung führt dann u. U. zu einem geringeren Schichtengewinn.

1) Auf eine Betrachtung der Nichtbeantwortung wird verzichtet.

2) Die Auswirkungen dieser fehlerhaften Schichtung auf die Schätzung der Schichtenmittelwerte ist jedoch zu beachten, wenn Informationen aus den Teilgesamtheiten von Interesse sind.

Systematische Fehler in der Auswahlgrundlage, z. B. wenn Elemente der Zielgesamtheit nicht aufgeführt sind, verzerren die Schätzung bei einfacher Zufallsauswahl und geschichteter Auswahl gleichermaßen. Beide Verfahren liefern eine erwartungstreue Schätzung bezogen auf die verzerrte Gesamtheit der in der Auswahlgrundlage aufgelisteten Untersuchungseinheiten.

Die Auswahl der Stichprobe erfolgt bei den betrachteten Verfahren unterschiedlich: Einmal wird die Gesamtheit als Ganzes betrachtet, während im anderen Fall die Auswahl in den Schichten getrennt erfolgt. Bei verzerrten Auswahlvorgängen können die auftretenden Verzerrungen der Schätzer sehr unterschiedlich sein. Da es sich bei den Schichten um kleinere Teilgesamtheiten handelt, sind diese leichter zu überblicken und zu kontrollieren. Tendenziell sind daher beim Schichtenauswahlverfahren geringere Verzerrungen zu erwarten. Andererseits ist die Gefahr, daß eine Verzerrung des Auswahlvorganges besteht, bei mehreren Auswahlvorgängen, wie es beim Schichtenauswahlverfahren der Fall ist, größer.

Auf die Wahrscheinlichkeit des Konfidenzintervalls wirkt, wie in Abschnitt 3.2.4 abgeleitet, die relative Verzerrung b^* bzw. b_g^* , der Quotient aus der absoluten Verzerrung und der Standardabweichung des Schätzers des jeweiligen Verfahrens. Der Vergleich der relativen Verzerrungen hängt wesentlich davon ab, ob die Schichtung dazu benutzt wird, die Varianz des Schätzers, d. h. die Länge des Konfidenzintervalls, zu reduzieren, oder ob die Schätzung mit niedrigerem Stichprobenumfang erfolgt. Auch Kombinationen, d. h. etwas geringerer Stichprobenumfang bei gleichzeitiger Verringerung der Varianz des Schätzers, sind möglich. Entscheidend bleibt die Qualität der Schichtung.

Wenn beim Schichtenauswahlverfahren eine Schätzung mit gleicher Varianz wie bei einer Schätzung mit einfacher Zufallsauswahl vorliegt, so kann wegen des geringeren Stichprobenumfangs mehr Sorgfalt auf den Auswahlvorgang verwendet werden. Es kann daher $b_g^* < b^*$ gelten, was zur Folge hat, das die Konfidenzaussage bei geschichteter Auswahl weniger durch die systematischen Fehler beeinflußt ist.

Werden die Vorteile des Schichtenauswahlverfahrens dazu benutzt, sowohl den Stichprobenumfang als auch die Varianz des Schätzers zu reduzieren, so ist keine eindeutige Aussage über die Verzerrung b_g^* der geschichteten Auswahl im Vergleich zu b^* der Referenzmethode möglich. Wegen des geringeren Stichprobenumfangs ist eine geringere absolute Verzerrung denkbar, andererseits wirkt die Varianzverringerung erhöhend auf die relative Verzerrung b_g^* . Bei gleichen relativen Verzerrungen ist der Einfluß auf die tatsächliche Wahrscheinlichkeit, mit der das Konfidenzintervall den zu schätzenden Parameter überdeckt, bei beiden Verfahren gleich, allerdings ist das Intervall bei einfacher Zufallsauswahl länger. Die geschichtete Auswahl weist eine präzisere Schätzung auf.

Der Einsatz des Schichtenauswahlverfahrens kann auch ausschließlich dazu benutzt werden, um die Varianz des Schätzers zu senken. In dieser Situation führen gleiche absolute Verzerrungen bei beiden Verfahren zu einer größeren relativen Verzerrung beim Schichtenauswahlverfahren. Damit reagiert die geschichtete Auswahl stärker auf die systematischen Fehler. Im folgenden Abschnitt soll versucht werden, diesen Effekt quantitativ zu erfassen.

3.4.2.2 Ein quantitativer Vergleich der Auswirkungen der Verzerrung auf die Schätzaussage bei einfacher und geschichteter Auswahl

Gehen wir davon aus, daß die Verfahren der einfachen und geschichteten Zufallsauswahl jeweils mit gleichem Gesamtstichprobenumfang durchgeführt werden, so ist es sinnvoll, für beide Verfahren gleiche absolute Verzerrungen anzunehmen, d. h. $b_{\bar{X}} = b_{\bar{X}_g}$. Der Einfluss der Schichtung auf die relative Verzerrung läßt sich über den Schichteneffekt darstellen. Durch Einsetzen von (3.67) und (3.68) in die Definition der relativen Verzerrung b_g^* folgt

$$b_g^* = \frac{b_{\bar{X}_g}}{\sqrt{\text{Var}(\bar{X}_g)}} = \frac{b_{\bar{X}_g}}{\sqrt{\text{Var}(\bar{X}) \cdot d^2}} = b^* \cdot \frac{1}{d}. \quad (3.70)$$

Bei positivem Schichteneffekt ist $\frac{1}{d} > 1$, so daß für das Schichtenauswahlverfahren eine größere relative Verzerrung besteht. Der negative Einfluß einer Verzerrung $b_{\bar{X}} = b_{\bar{X}_g}$ auf die Schätzaussage ist daher beim Schichtenauswahlverfahren stärker.

Eine Möglichkeit, die Auswirkungen einer Verzerrung auf die Konfidenzaussage bei einfacher und geschichteter Auswahl quantitativ zu vergleichen, besteht darin, die Veränderung der Wahrscheinlichkeitsaussage des Konfidenzintervalls bei geschichtetem Schätzer für verschiedene Schichteneffekte und relative Verzerrungen b^* zu berechnen.

Durch Einsetzen von (3.70) in (3.60) läßt sich die tatsächliche Wahrscheinlichkeit γ_g' ermitteln, mit der das Konfidenzintervall (3.59) den wahren zu schätzenden Parameter überdeckt

$$\begin{aligned} \gamma_g' &= F(0;1) \left(z_{\alpha/2} - b^* \frac{1}{d} \right) \\ &\quad - F(0;1) \left(-z_{\alpha/2} - b^* \frac{1}{d} \right). \end{aligned} \quad (3.71)$$

Die Wahrscheinlichkeit γ'_g hängt in dieser Darstellung vom Konfidenzniveau γ , der relativen Verzerrung b^* des Schätzers bei einfacher Zufallsauswahl und dem Faktor $\frac{1}{d}$ ab.

Entscheidend für den Einfluß der Schichtung auf die tatsächliche Wahrscheinlichkeit γ'_g ist (vgl. (3.68) und (3.69)) der Faktor

$$\frac{1}{d} = \left[1 - \frac{D}{\text{Var}(\bar{X})} \right]^{-\frac{1}{2}} = (1-D^*)^{-\frac{1}{2}},$$

im wesentlichen also der relative Schichteneffekt D^* , das Verhältnis zwischen dem Schichteneffekt D und der Varianz des Schätzers bei einfacher Zufallsauswahl.

In den Tabellen 3.6a bis 3.6d (vgl. S. 109 - 112) ist γ'_g ¹⁾ in Abhängigkeit von D^* und b^* für verschiedene Konfidenzniveaus angegeben. Gelingt es nicht, durch die Schichtung die Varianz des Schätzers $\bar{X}_g$ zu verbessern, so läßt sich Tabelle 3.2 anwenden; die Werte von γ'_g für $D = 0$ sind deshalb in den Tabellen 3.6a bis 3.6d nicht mit aufgeführt.

Zusätzlich zu γ'_g ist in Tabelle 3.6a für die verschiedenen D^* der Faktor $\frac{1}{d}$ mit aufgeführt. Durch Multiplikation von b^* mit $\frac{1}{d}$ erhält man nach (3.70) die relative Verzerrung $b_{\bar{X}_g}^*$ des geschichteten Schätzers. Ermöglicht die Schichtung beispielsweise ein $D^* = 0,5$, so beträgt die relative Verzerrung des geschichteten Schätzers $b_{\bar{X}_g}^* = \sqrt{2} \cdot b^*$. Die Wahrscheinlichkeit γ'_g sinkt für ein Konfidenzniveau $\gamma = 0,95$ auf 0,9135.

In Schaubild 3.4 ist die Abhängigkeit zwischen γ'_g und der relativen Verzerrung b^* für ausgewählte D^* und dem Konfidenzniveau $\gamma = 0,95$ dargestellt. Bei

1) Wegen der numerischen Bedeutung von γ'_g sind in den Werten Rundungsfehler enthalten. Die ersten 4 Stellen hinter dem Komma sind genau.

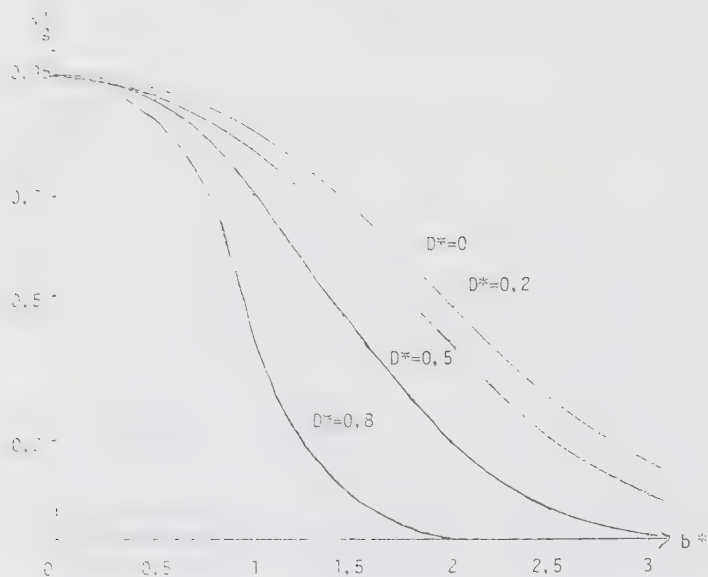


Schaubild 3.4: Die Abhängigkeit zwischen γ'_G und der relativen Verzerrung des Schätzers bei einfacher Zufallsauswahl für ausgewählte D^* bei einem Konfidenzniveau $\alpha = 0,95$.

instiger Schichtung (z.B. $D^* = 0,3$) sinkt die tatsächliche Wahrscheinlichkeit, mit der die Konfidenzaussage zutrifft, mit steigender Verzerrung schnell.

Bearbeitet man die Tabelle 3.6 zeilenweise, so wird der Einfluß der Schichtung auf die Auswirkungen der Verzerrung deutlich.

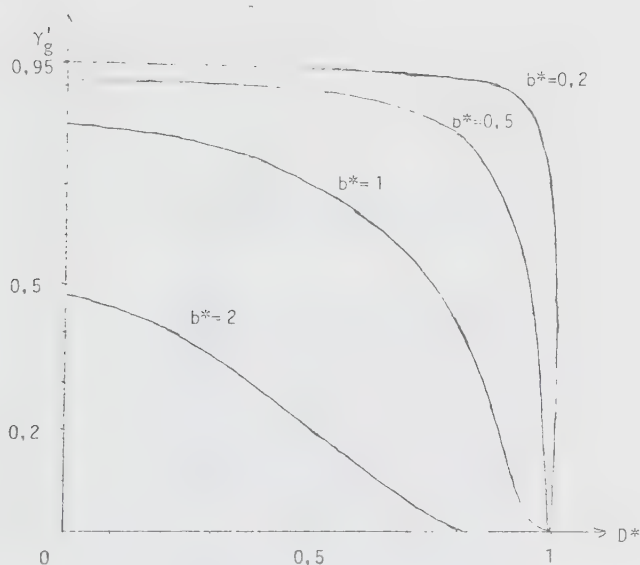


Schaubild 3.5: Die Abhängigkeit zwischen γ'_g und relativem Schichteneffekt D^* für ausgewählte b^* bei einem Konfidenzniveau $\gamma = 0,95$.

Für eine Schätzung bei einfacher Zufallsauswahl mit $\gamma = 0,95$ beträgt die tatsächliche Wahrscheinlichkeit, mit der das Konfidenzintervall den Parameter μ überdeckt, bei einer relativen Verzerrung von $b^* = 0,5$ noch etwa 0,92. Kann die Varianz des Mittelwertschätzers durch die Schichtung um 50% reduziert werden ($D^* = 0,5$), so sinkt γ'_g auf rund 0,89.

Die Abhängigkeit zwischen γ'_g und D^* ist in Schaubild 3.5 für verschiedene relative Verzerrungen bei einem Konfidenzniveau $\gamma = 0,95$ dargestellt.

Es zeigt sich, daß bei geringer relativer Verzerrung (z. B. $b^* = 0,2$) die Kurve bis zu einem relativen Schichteneffekt $D^* = 0,9$ sehr flach verläuft; γ'_g beträgt noch etwa 0,90. Erst bei noch weiter steigendem Schichteneffekt sinkt γ'_g schnell. Die Schätzungsaussage des Schichtenauswahlverfahrens reagiert auf geringe Verzerrungen nur bei sehr starkem Schichtengewinn empfindlich.

Mit zunehmender Verzerrung der Schätzung ist der Kurvenverlauf bereits für kleinere D^* steiler, d. h. der Schichtengewinn führt zu einer deutlichen Steigerung der Auswirkungen der Verzerrung auf die Schätzungsaussage. Bei der Interpretation der Ergebnisse ist jedoch zu beachten, daß die Grenzen des Konfidenzintervalls durch die Standardabweichung des geschätzten Schätzers bestimmt sind. Das Konfidenzintervall ist um den Faktor $d < 1$ kürzer als ein entsprechendes Intervall für die einfache Zufallsauswahl.

b*	D*	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.00		.6826158	.68269158	.68269158	.68269158	.68269158	.68269158	.68269158	.68269158	.68269158
.05		.682619744	.68269158	.68269158	.68269158	.68269158	.68269158	.68269158	.68269158	.68269158
.1		.68269766	.68269158	.68269158	.68269158	.68269158	.68269158	.68269158	.68269158	.68269158
.15		.676657402	.675917307	.675917307	.675917307	.675917307	.675917307	.675917307	.675917307	.675917307
.2		.672016943	.67069326	.66989955	.667674101	.671894105	.666297499	.664760778	.659774553	.650740356
.25		.666001058	.664031671	.661408512	.657919149	.653007304	.645932686	.637974592	.630976566	.624025723
.3		.658873584	.655974553	.652387931	.647295976	.640419065	.630244035	.615392701	.600976566	.584779816
.4		.64972637	.638078303	.629442992	.620958075	.609263031	.592103102	.565518215	.512559732	.494797102
.5		.618503706	.616892666	.6017356	.58506598	.571200193	.546263723	.505835754	.432924986	.27649954
.6		.572103102	.591612715	.568083112	.551106816	.527353114	.494797408	.443995137	.356710772	.182919283
.75		.546253723	.531148531	.512796605	.49914664	.485147541	.411846505	.347056631	.245172393	.0847802555
1		.45844771	.435934487	.40633745	.37455757	.331473059	.275149054	.202115031	.107603792	.015202503
1.5		.27549054	.245472383	.211321468	.172049782	.13074607	.0847027554	.04980278	.27681946	9.305197e-05
2		.137940791	.07609783	.0314453027	.0546548177	.0335976081	.0153827802	4.0031701E-03	2.5913123E-04	5.0528981E-08
2.5		.052092702	.032456979	.0233681207	.0129452766	5.6107694E-03	1.57386180E-03	1.8352265E-04	2.2142012E-06	2.49742244E-12
3		.0152092702	9.2760335E-03	4.05685935E-03	2.0324511E-03	5.92038079E-04	7.0757194E-05	3.70809155E-06	5.7079230E-07	1.0614031E-17

Tabelle 3.6a: Die tatsächliche Wahrscheinlichkeit γ'_g in Abhängigkeit von b^* und D^* für ein $z_{\alpha/2} = 1$, Konfidenzniveau $\gamma = 0,6826$.

b*	D*	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.00		.900074100	.900034109	.900034109	.900034109	.900034100	.900034100	.900034109	.900034109	.900034109
.03		.897562718	.897594071	.897428303	.897427438	.899106123	.898874157	.898620214	.897914537	.897776271
.1		.898149378	.897914537	.89761105	.897208308	.89664343	.895796271	.894380876	.891561070	.893116061
.15		.895786291	.895366939	.894586472	.893673708	.892409719	.890726472	.887337847	.881007790	.872181095
.2		.895573937	.89235635	.89235635	.890973369	.886492395	.883116051	.877497424	.866793978	.852785153
.25		.890235043	.894090119	.894974341	.892413138	.887901856	.883641037	.876166053	.867393041	.854775573
.3		.891516061	.891008778	.891029745	.891467240	.882101986	.873641037	.86166083	.844775573	.821379806
.4		.870023274	.866293978	.861507005	.855133954	.846251097	.832895153	.811963233	.787764179	.761251471
.5		.853294238	.847490641	.840081914	.830373171	.816520117	.796135981	.762692381	.719830472	.670079953
.6		.832705163	.824735804	.81417913	.800194468	.780900352	.752139986	.705622331	.647789155	.580177953
.65		.796135804	.785314112	.767427305	.746237574	.717090315	.674616279	.607721251	.495760871	.373670362
.7		.719234765	.698030472	.671310921	.636617976	.590131247	.524932673	.429024372	.271793243	.130532979
1		.574032573	.466160872	.440939617	.395165646	.316038386	.236703662	.130532979	.607721251	.373670362
1.5		.331526721	.227103244	.227967923	.174365599	.118313483	.064935365	.0224010735	2.34005015-01	9.7217413E-04
2		.161018477	.12504729	.0896601599	.0567667405	.0293417679	.0105030706	1.75353294E-03	3.986432E-05	1.9151583E-10
2.5		.064595395	.0437139699	.021067928	.0129399622	.69273253E-03	9.72674731E-04	6.33035012E-05	7.0757253E-07	2.7103170E-15

Tabelle 3.6b: Die tatsächliche Wahrscheinlichkeit γ'_g
in Abhängigkeit von b^* und D^* für ein
 $\alpha/2 = 1,645$, Konfidenzniveau $\gamma \approx 0,9$.

b*	D*	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.00	.950007977	.950007977	.950007977	.950007977	.950007977	.950007977	.950007977	.950007977	.950007977	.950007977
.05	.949487358	.949487358	.949487358	.949487358	.949487358	.949487358	.949487358	.949487358	.949487358	.949487358
.1	.94873429	.94873429	.94873429	.94873429	.94873429	.94873429	.94873429	.94873429	.94873429	.94873429
.15	.947839429	.947839429	.947839429	.947839429	.947839429	.947839429	.947839429	.947839429	.947839429	.947839429
.2	.946801559	.946801559	.946801559	.946801559	.946801559	.946801559	.946801559	.946801559	.946801559	.946801559
.25	.945620589	.945620589	.945620589	.945620589	.945620589	.945620589	.945620589	.945620589	.945620589	.945620589
.3	.94431159	.94431159	.94431159	.94431159	.94431159	.94431159	.94431159	.94431159	.94431159	.94431159
.35	.942876886	.942876886	.942876886	.942876886	.942876886	.942876886	.942876886	.942876886	.942876886	.942876886
.4	.9413114	.9413114	.9413114	.9413114	.9413114	.9413114	.9413114	.9413114	.9413114	.9413114
.45	.939629676	.939629676	.939629676	.939629676	.939629676	.939629676	.939629676	.939629676	.939629676	.939629676
.5	.93782013	.93782013	.93782013	.93782013	.93782013	.93782013	.93782013	.93782013	.93782013	.93782013
.55	.935874576	.935874576	.935874576	.935874576	.935874576	.935874576	.935874576	.935874576	.935874576	.935874576
.6	.933798474	.933798474	.933798474	.933798474	.933798474	.933798474	.933798474	.933798474	.933798474	.933798474
.65	.931598684	.931598684	.931598684	.931598684	.931598684	.931598684	.931598684	.931598684	.931598684	.931598684
.7	.929282961	.929282961	.929282961	.929282961	.929282961	.929282961	.929282961	.929282961	.929282961	.929282961
.75	.926854474	.926854474	.926854474	.926854474	.926854474	.926854474	.926854474	.926854474	.926854474	.926854474
.8	.92431361	.92431361	.92431361	.92431361	.92431361	.92431361	.92431361	.92431361	.92431361	.92431361
.85	.921673861	.921673861	.921673861	.921673861	.921673861	.921673861	.921673861	.921673861	.921673861	.921673861
.9	.918943761	.918943761	.918943761	.918943761	.918943761	.918943761	.918943761	.918943761	.918943761	.918943761
1	.916122961	.916122961	.916122961	.916122961	.916122961	.916122961	.916122961	.916122961	.916122961	.916122961
1.5	.847407689	.847407689	.847407689	.847407689	.847407689	.847407689	.847407689	.847407689	.847407689	.847407689
2	.647407689	.647407689	.647407689	.647407689	.647407689	.647407689	.647407689	.647407689	.647407689	.647407689
2.5	.447407689	.447407689	.447407689	.447407689	.447407689	.447407689	.447407689	.447407689	.447407689	.447407689
3	.247407689	.247407689	.247407689	.247407689	.247407689	.247407689	.247407689	.247407689	.247407689	.247407689
3.5	.047407689	.047407689	.047407689	.047407689	.047407689	.047407689	.047407689	.047407689	.047407689	.047407689

Tabelle 3.6c: Die tatsächliche Wahrscheinlichkeit γ' in Abhängigkeit von b^* und D^* für ein $z_{\alpha/2} = 1,96$, Konfidenzniveau $\gamma = 0,95$.

b*	D*	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.00	.990122119	.970122119	.990122119	.990122119	.990122119	.990122119	.990122119	.990122119	.990122119	.990122119
.05	.99019523	.970066695	.990190178	.990190178	.990190178	.990190178	.990190178	.990190178	.990190178	.990190178
.1	.989710695	.969559871	.989532534	.989532534	.989532534	.989532534	.989532534	.989532534	.989532534	.989532534
.15	.989192517	.968975331	.988974344	.988974344	.988974344	.988974344	.988974344	.988974344	.988974344	.988974344
.2	.988495016	.968248925	.987976767	.987976767	.987976767	.987976767	.987976767	.987976767	.987976767	.987976767
.25	.987505397	.967170754	.986758049	.986758049	.986758049	.986758049	.986758049	.986758049	.986758049	.986758049
.3	.986320064	.965027218	.985027218	.985027218	.985027218	.985027218	.985027218	.985027218	.985027218	.985027218
.4	.983210124	.96229765	.98110057	.98110057	.98110057	.98110057	.98110057	.98110057	.98110057	.98110057
.5	.979919772	.957514915	.975542897	.975542897	.975542897	.975542897	.975542897	.975542897	.975542897	.975542897
.6	.973610766	.951100893	.96824734	.96824734	.96824734	.96824734	.96824734	.96824734	.96824734	.96824734
.75	.96285408	.950889413	.953619055	.953619055	.953619055	.953619055	.953619055	.953619055	.953619055	.953619055
1	.97349494	.928021797	.91684681	.901254058	.878123708	.841007997	.784534225	.704534225	.604534225	.484534225
1.5	.841859997	.816719903	.78445367	.740855172	.6767112	.582501373	.453082252	.289102439	.141963347	.023051019
2	.684473911	.634554225	.573168883	.499264456	.40079335	.289102439	.141963347	.023051019	1.395199566	5.021321E-04
2.5	.417776425	.41464823	.34168469	.258652919	.16847219	.091974078	.023051019	1.395199566	5.021321E-04	2.474656E-12
3	.200184437	.219428939	.157276483	.076022719	.048107568	.0152522086	1.891912E-03	1.9271507E-05	2.474656E-12	

Tabelle 3.6d: Die tatsächliche Wahrscheinlichkeit γ'_g
in Abhängigkeit von b^* und D^* für ein
 $z_{\alpha/2} = 2.58$, Konfidenzniveau $\gamma = 0.99$.

4. Die Mittelwertschätzung unter Berücksichtigung systematischer Fehler in den Antworten

4.1 Abgrenzungen

In Kapitel 2 dieser Arbeit wurde die Trennung zwischen systematischen Fehlern, die bei korrekten Antworten der Untersuchungseinheiten auftreten und solchen, die die individuellen Daten beeinflussen, vorgenommen. Erstere, sie erzeugen eine konstante Verzerrung des Schätzers, sind in Kapitel drei abgehandelt. In diesem Abschnitt sollen die Einflüsse von Fehlern in den Daten auf die Schätzung untersucht werden.

Die auftretenden Fehler hängen entscheidend von der konkreten Erhebungssituation ab. Je nach Erhebungssystem werden die Fehler als Meß- oder Antwortfehler bezeichnet. In dieser Arbeit sind die Betrachtungen mehr auf eine Befragung ausgerichtet.

Seit dem zweiten Weltkrieg sind mehrere Modelle zur Stichprobentheorie entwickelt worden, in denen versucht wurde, Fehler in den Daten zu berücksichtigen¹⁾. Sie unterscheiden sich durch die Annahmen bezüglich des Untersuchungsmerkmals und in den zugrundeliegenden Fehlern. Einige Autoren berücksichtigen Antwortfehler allgemein, ohne auf spezielle Ursachen einzugehen, andere wiederum beziehen ihre Ausführungen auf spezielle Fehlerursachen wie z. B. Interviewereinfluß. Des weiteren werden zum Teil Abhängigkeiten

1) Vgl.: z. B. Hansen u. a., 51, S. 155 ff; Strecker, 71, S. 141 ff; Strecker, 80, S. 390 ff; Stenger, 77, S. 391 ff; Cochran, 72, S. 437 ff; Sukhatme u. Sukhatme, 70, S. 390 ff; Pokropp, 80, S. 201 ff; Krug u. Nourney, 82, S. 194 ff.

zwischen den Fehlern verschiedener Untersuchungseinheiten oder zwischen Antwortverzerrungen und wahren Merkmalsausprägung der betrachteten Einheiten berücksichtigt. Gerade diese Abhängigkeiten erschweren die Beschreibung der Fehler in den Modellen.

In einigen Arbeiten sind nicht nur die Antwortfehler in den Stichprobenverfahren berücksichtigt, es sind auch Möglichkeiten zur Schätzung der durch die Fehler verursachten Verzerrung mit einbezogen. In diesem Zusammenhang sind die Wiederholungszählung¹⁾ und das Modell der 'interpenetrating subsamples', es wurde bereits 1946 von Mahalanobis²⁾ entwickelt, hervorzuheben.

Ziel dieser Arbeit ist es nicht, die einzelnen Modelle mit ihren Besonderheiten vorzustellen. Vielmehr soll versucht werden, exemplarisch die Struktur der Modelle aufzuzeigen und die Auswirkungen der Fehler in den Daten auf die Ergebnisse der Stichprobe darzustellen. Auf Möglichkeiten und Schwierigkeiten bei der Schätzung der Verzerrung und der Varianz des Stichprobenschätzers und seiner Komponenten, wie sie in einigen Modellen behandelt sind, wird nur am Rande eingegangen.

Wie bereits im Kapitel 3 ist bei den Betrachtungen ein quantitatives Merkmal X unterstellt. Die Schätzung des Mittelwertes μ der Merkmalsausprägungen x_j der Elemente aus der betrachteten Gesamtheit erfolgt, wie bisher, aus der Stichprobe $\underline{U}$.

Im Unterschied zu Kapitel 3 sei im folgenden unterstellt, daß die Auswahlgrundlage und der Auswahlvor-

1) Vgl.: z. B. Strecker, 80, S. 390 ff, Strecker u. Wiegert, 83, S. 11 ff; Stenger, 77, S. 394 ff;

2) Vgl.: Mahalanobis, 46, S. 325 ff; vgl. auch Stenger, 77, S. 397 ff; Cochran, 72, S. 447 ff.

gang - bei einfacher und geschichteter Auswahl - nicht durch systematische Fehler beeinträchtigt sind. Fehler treten nur in den individuellen Angaben auf.

Die Überlegungen können an die Ausführungen aus Kapitel 2 anknüpfen, wo die Arten und Ursachen der Fehler in den Antworten dargestellt sind und auch bereits ein Ansatz zur Berücksichtigung der Fehler in einem Fehlerterm ϵ'_j skizziert ist.

Die Ursachen der systematischen Fehler in den Antworten der einzelnen Untersuchungseinheiten aus G legen die Vermutung nahe, daß die bei der Untersuchung ermittelte Merkmalsausprägung der Einheit u_j nicht unbedingt mit dem wahren Wert x_j der Merkmalsausprägung übereinstimmt. In den Antworten sind die x_j durch einen zufälligen Fehlerterm ϵ'_j überlagert. Wegen der Antwortfehler ist die Antwort jedes betrachteten Elements aus G - auch ohne die Zufallsauswahl beim Ziehen der Stichprobe - Realisation einer Zufallsvariablen

$$Y'_j = x_j + \epsilon'_j,$$

die das Antwortverhalten der Individuen (vgl. (2.12)) beschreibt.

Wegen der Annahme eines quantitativen Merkmals ist es notwendig, die Realisationen der Fehlervariablen ϵ'_j ebenfalls als quantitativ anzunehmen. Die Ausführungen beschränken sich im folgenden auf Antwortfehler, die in der befragten Einheit begründet sind, d. h. es wird von einem Interviewereinfluß abgesehen. Die Ergebnisse lassen sich auch auf den Interviewereinfluß übertragen, wobei die verschiedenen Interviewer durch einen zusätzlichen Index einbezogen werden müssen¹⁾.

1) Ein Modell mit Interviewereinfluß stellt z. B. Hansen u. a., 51, S. 155 ff und Sukhatme u. Sukhatme, 70, S. 390 ff, vor.

Um die Ausführungen nicht unnötig einzuschränken, sind keine Annahmen über die Verteilung der ε_j^i unterstellt, lediglich die Existenz der Erwartungswerte

$$E(\varepsilon_j^i) = b_j, \quad b_j \in \mathbb{R}, \quad j = 1, 2, \dots, N,$$

und der Varianzen

$$\text{Var}(\varepsilon_j^i) = \sigma_j^2, \quad j = 1, 2, \dots, N,$$

sei sichergestellt. Sowohl die b_j als auch die σ_j^2 können für verschiedene Elemente aus der Gesamtheit G verschiedene Werte annehmen.

In G läßt sich dann die durchschnittliche Antwortverzerrung

$$\bar{b} = \frac{1}{N} \sum_j b_j, \quad (4.1)$$

die Varianz der individuellen Antwortverzerrungen

$$\sigma_b^2 = \frac{1}{N} \sum_j (b_j - \bar{b})^2 \quad (4.2)$$

und die durchschnittliche Antwortvarianz

$$\bar{\sigma}^2 = \frac{1}{N} \sum_j \sigma_j^2 \quad (4.3)$$

definieren.

Das Pendant zur individuellen Verzerrung b_j ist die Zufallsvariable

$$v_j^i = \varepsilon_j^i - b_j, \quad (4.4)$$

der Zufallsfehler der individuellen Antworten. Für v_j^i wird auch die Bezeichnung Antwortvariabilität be-

nutzt¹⁾. Die Varianz σ_j^2 , die man auch als Antwortvarianz der Einheit u_j bezeichnen kann, ist ein Maß für den Zufallsfehler $V_j^!$.

Die Parameter der Zufallsvariablen $Y_j^!$ lauten:

$$E(Y_j^!) = x_j + b_j$$

$$\text{Var}(Y_j^!) = \text{Var}(\epsilon_j^!) = \sigma_j^2.$$

Im folgenden soll noch gelten, daß die Antwortfehler der Elemente aus G unabhängig voneinander sind. Wahrscheinlichkeitstheoretisch bedeutet das: Die $\epsilon_j^!$ sind für $j = 1, 2, \dots, N$ unabhängige Zufallsvariablen. Diese Einschränkung ist durchaus sinnvoll, denn die Antworten der bereits befragten Einheiten sind den noch zu befragenden Einheiten nicht bekannt. Das heißt jedoch nicht, daß die Antworten der u_j nicht alle in die gleiche Richtung verzerrt sein können (z. B. $b_j > 0$ für $j = 1, 2, \dots, N$), lediglich wird vermieden, daß die Abweichungen der Antwortfehler eines Elements vom Erwartungswert b_j von den Antworten anderer Einheiten abhängt.

Eine solche Abhängigkeit kann unter Umständen durch einen Interviewer entstehen, wenn dieser unbewußt einen Ausgleich der Antworten erzeugt. Beeinflusst der Interviewer die Antworten immer in die gleiche Richtung, so ist auch eine andere Interpretation möglich: Der Interviewer erzeugt zusätzlich einen Antwortfehler ϵ_j^* mit $E(\epsilon_j^*) = b(j)$ und $\text{Var}(\epsilon_j^*)$. Eine Abhängigkeit der $\epsilon_j^!$, $j = 1, 2, \dots, N$, ist dann nicht gegeben. Oben wurde bereits darauf hingewiesen, daß der Interviewereinfluß in den Betrachtungen nicht berücksichtigt wird. Daher ist die Annahme der Unabhängigkeit der Antwortfehler in dieser Arbeit sinnvoll.

1) Vgl.: Pokropp, 30, S. 203.

Auch der Einfluß der Fragebogengestaltung auf die Antworten führt nicht unbedingt zu einer Abhängigkeit der ϵ'_j . Durch das Erhebungssystem, auf das auch die Fragebogengestaltung Einfluß hat, wird die Lage und die Art der Verteilung der ϵ'_j bestimmt. Verschieden aufgebaute Fragebögen zum gleichen Erhebungsgegenstand führen meist zu anderen b_j bzw. σ_j^2 , eine Abhängigkeit der Antworten entsteht durch die Fragebogengestaltung nicht notwendigerweise.

Zielsetzung dieses Kapitels ist es, unter Zugrundelegung der bisher getroffenen Annahmen, die Fehler in den Antworten bei einfacher Zufallsauswahl und beim Schichtenauswahlverfahren zu beschreiben und die Auswirkungen auf die jeweilige Schätzaussage zu untersuchen. Dabei finden die in Kapitel 2 erwähnten Sonderfälle:

- (1) verzerrungsfreie, aber variable Antworten, d. h. $b_j = 0$ und $\sigma_j^2 > 0$;
- (2) verzerrte Antworten, die aber nicht variabel sind, d. h. $b_j \neq 0$ und $\sigma_j^2 = 0$;

Berücksichtigung. Mit $b_j = 0$ und $\sigma_j^2 = 0$ für alle Elemente aus G entsteht die einfache Zufallsauswahl ohne Antwortfehler.

4.2 Die einfache Zufallsauswahl bei Antwortfehlern

4.2.1 Die Stichprobe bei Antwortfehlern

Den Ausführungen in diesem Abschnitt liegt das Modell der einfachen Zufallsauswahl zugrunde, wie es bereits in Abschnitt 3.2.1 beschrieben ist.

Aus der Gesamtheit G wird zur Schätzung von μ eine Stichprobe $\underline{U}$ vom Umfang n mit bzw. ohne Zurücklegen gezogen. Der Auswahlvorgang der Stichprobe bleibt unbeeinflusst von den Antwortfehlern, d. h. die Menge der Realisationen der U_i aus $\underline{U}$ ist G , wobei für alle u_j die gleiche Wahrscheinlichkeit für das Eintreten gilt.

Erst bei der Befragung der ausgewählten Einheiten spielt das Antwortverhalten, wie es in Abschnitt 4.1 skizziert ist, eine Rolle. Den U_i ist nicht mehr ein X_i zugeordnet, sondern mit (2.17) die Zufallsvariable Y_i : Wenn U_i die Realisation u_j annimmt, gilt für Y_i die Realisation Y'_j .

Die Folge der Realisation von Y_i ist daher:

$$\begin{aligned} G_Y &= (Y'_j / j = 1, 2, \dots, N) \\ &= (x_j + \varepsilon'_j / j = 1, 2, \dots, N). \end{aligned}$$

Wegen der direkten Zuordnung zwischen u_j und Y'_j treten die Realisationen aus G_Y alle mit der gleichen Wahrscheinlichkeit

$$P(Y_i = Y'_j) = \frac{1}{N}, \quad j = 1, 2, \dots, N, \quad (4.5)$$

ein. Aus der Stichprobe $\underline{U}$ entsteht unter Berücksichtigung der Befragung die Stichprobe

$$\underline{Y} = (Y_i / i = 1, 2, \dots, n). \quad (4.6)$$

In $\underline{Y}$ kommen sowohl der Auswahlvorgang unter der Untersuchungseinheiten als auch die zufälligen Antwortfehler der ausgewählten Einheiten zum Tragen.

Die additive Überlagerung der wahren Merkmalsausprägungen der ausgewählten Untersuchungseinheit mit dem Antwortfehler entsprechend (2.12) erlaubt, unter Beachtung der direkten Zuordnung zwischen Y_j' und u_j , eine Zerlegung der Zufallsvariablen Y_i :

$$Y_i = X_i + \varepsilon_i, \quad i = 1, 2, \dots, n. \quad (4.7)$$

Mit (2.9) ist die Zufallsvariable X_i bereits definiert als die durch die einfache Zufallsauswahl entstehende Auswahl unter den Merkmalsausprägungen der Einheiten aus G .

Die Antwortfehler ε_j' sind in der Zufallsvariablen ε_i erfasst, überlagert vom Zufallsprozeß 'Stichproben Ziehen', denn (4.7) schließt auch die Auswahl unter den ε_j' mit ein. Beim Auswahlvorgang der Y_j' aus G_Y sind die ε_j' die Realisationen der Zufallsvariablen ε_i . Erst bei der anschließenden Befragung nehmen die Zufallsvariablen ε_j' jeweils Realisationen an.

Gleichung (4.7) ermöglicht die Aufspaltung des Zufallsvektors $\underline{Y}$ in

$$\begin{aligned} \underline{Y} &= (X_i + \varepsilon_i / i = 1, 2, \dots, n) \\ &= (X_i / i = 1, 2, \dots, n) + (\varepsilon_i / i = 1, 2, \dots, n) \\ &= \underline{X} + \underline{\varepsilon}. \end{aligned} \quad (4.8)$$

Es erleichtert die spätere Beweisführung, wenn wir folgende bedingten Erwartungswerte und Varianzen betrachten. Als Bedingung sei eine bestimmte Stichpro-

be $\underline{U}$ aus G bzw. ein bestimmtes Element u_j aus G gegeben. Die Erwartungswertbildung erfolgt dann über die Antworten der Elemente. Es gilt:

$$E(\epsilon_i | U_i = u_j) = E(\epsilon_j^*) = b_j \quad (4.9)$$

$$\begin{aligned} \text{Var}(\epsilon_i | U_i = u_j) &= \text{Var}(\epsilon_i | u_j - \mathbb{I}(\epsilon_i | u_j))^2 \\ &= \text{Var}(\epsilon_j^*) = \sigma_j^2 \end{aligned} \quad (4.10)$$

$$E(\epsilon_i | \underline{U}) = E(\epsilon_i | U_i = u_j) = b_j \quad (4.11)$$

$$\text{Var}(\epsilon_i | \underline{U}) = \text{Var}(\epsilon_i | U_i = u_j) = \sigma_j^2, \quad (4.12)$$

falls u_j in $\underline{U}$ enthalten ist.

Der Übergang von der Bedingung $\underline{U}$ auf u_j ist erlaubt, denn nach den oben getroffenen Annahmen unabhängiger Antworten der Individuen, insbesondere der Antwortfehler haben die anderen in $\underline{U}$ ausgewählten Elemente keine Auswirkungen. Es muß lediglich vorausgesetzt werden, daß u_j in der Stichprobe enthalten ist.

Ist $\underline{U}$ mit Zurücklegen gezogen, so kann ein Element u_j mehrmals in die Stichprobe gelangen. Die Beziehungen (4.11) und (4.12) setzen dann voraus, daß die jeweiligen Antworten der Einheit u_j unabhängig voneinander sind¹⁾.

Den Betrachtungen liegt das Modell der einfachen Zufallsauswahl zugrunde, d. h. die Y_i , $i = 1, 2, \dots, n$, sind identisch verteilt, wobei zu beachten ist, daß beim Ziehen ohne Zurücklegen eine Abhängigkeit zwischen den Y_i besteht.

Der Erwartungswert von Y_i läßt sich mit (4.7), (4.9), (4.5), (3.6) und (4.1) ableiten:

1) Auf diese Annahme wird im Anschluß an die Ableitung der Varianz des Schätzers bei Antwortfehlern nochmals eingegangen.

$$\begin{aligned}
 E(Y_i) &= E(X_i) + E(\epsilon_i) \\
 &= E(X_i) + E(E(\epsilon_i | U_i = u_j)) \\
 &= \sum_{j=1}^N \frac{1}{N} x_j + \sum_{j=1}^N \frac{1}{N} b_j \\
 &= \mu + \bar{b}.
 \end{aligned} \tag{4.13}$$

Bei der Berechnung von $E(\epsilon_i)$ ist der Erwartungswert über alle Fehlerausprägungen der Einheit U_i und über den Auswahlvorgang zu bilden, d. h. alle Realisationen, die U_i annehmen kann, sind zu betrachten. Der Erwartungswert der Zufallsvariablen ϵ_i ist die durchschnittliche Antwortverzerrung der u_j in der Gesamtheit, also

$$E(\epsilon_i) = \bar{b}. \tag{4.14}$$

Da wir Unabhängigkeit zwischen Merkmalsausprägung und Fehler unterstellen, läßt sich die Varianz der Y_i zerlegen:

$$\text{Var}(Y_i) = \text{Var}(X_i) + \text{Var}(\epsilon_i). \tag{4.15}$$

Mit (3.7) ist bereits $\text{Var}(X_i)$ gegeben; die Berechnung der Varianz der ϵ_i muß den Zufallsfehler bei der Antwort des ausgewählten Elements und den Auswahlprozeß beachten. Dies ist über die Varianzaufspaltung¹⁾ möglich:

$$\begin{aligned}
 \text{Var}(\epsilon_i) &= E[\text{Var}(\epsilon_i | U_i)] + \text{Var}\{E(\epsilon_i | U_i)\}^2 \\
 &= E[\text{Var}(\epsilon_i | U_i)] + E[E(\epsilon_i | U_i) - E[E(\epsilon_i | U_i)]]^2.
 \end{aligned} \tag{4.16}$$

Innerhalb der eckigen Klammern wird der Antwortfehler des ausgewählten Individuums berücksichtigt. Unter Beachtung von (4.12), (4.9), (4.14), (4.1) und (4.2) folgt aus (4.16)

1) Vgl.: z. B. Pokropp, 80, S. 233 f.

$$\begin{aligned}\text{Var}(\varepsilon_i) &= \sum_{j=1}^N \frac{1}{N} \sigma_j^2 + \sum_{j=1}^N \frac{1}{N} (b_j - \bar{b})^2 \\ &= \bar{\sigma}^2 + \sigma_b^2.\end{aligned}\quad (4.17)$$

Der erste Summand ist das arithmetische Mittel $\bar{\sigma}^2$ der individuellen Antwortvarianzen der Elemente aus G. Zusätzlich bestimmen auch die Unterschiede der individuellen Verzerrungen, gemessen durch die Varianz σ_b^2 der Verzerrungen, die Varianz der Zufallsvariablen ε_i . Durch Einsetzen von (4.17) in (4.15) ergibt sich mit (3.7)

$$\text{Var}(Y_i) = \sigma^2 + \bar{\sigma}^2 + \sigma_b^2. \quad (4.18)$$

4.2.2 Die Mittelwertschätzung bei Antwortfehlern

4.2.2.1 Erwartungswert und Varianz des Mittelwertschätzers bei Antwortfehlern

Das Modell der einfachen Zufallsauswahl benutzt das arithmetische Mittel der Zufallsvariablen 'Antwort des Befragten' in der Stichprobe als Schätzer für μ . Entsprechend ist bei Antwortfehlern

$$\bar{Y} = \frac{1}{N} \sum_i Y_i \quad (4.19)$$

ein Schätzer für μ . Es ist im folgenden zu untersuchen, welche Auswirkungen die Antwortfehler auf die Eigenschaften des Schätzers $\bar{Y}$ haben. Zum einen ist der Erwartungswert von $\bar{Y}$ zu berechnen und zum anderen die Varianz des Schätzers zu bestimmen¹⁾.

1) Die Beschreibung der einfachen Zufallsauswahl bei Antwortfehlern basiert im wesentlichen auf Pokropp, 80, S. 196 ff und Cochran, 72, S. 437 ff. Die Beweisführung ist jedoch anders aufgebaut.

Mit (4.13) läßt sich der Erwartungswert von $\bar{Y}$ leicht angeben:

$$\begin{aligned} E(\bar{Y}) &= \frac{1}{n} \sum_i E(Y_i) \\ &= \frac{1}{n} \sum_i (\mu + \bar{b}) \\ &= \mu + \bar{b}. \end{aligned} \tag{4.20}$$

Durch die Antwortfehler ist der Schätzer um das arithmetische Mittel $\bar{b}$ der individuellen Verzerrungen b_j in der Gesamtheit verzerrt. Je nach der konkreten Erhebungssituation bzw. den Modellannahmen bezüglich der Y_j zur Beschreibung des Antwortverhaltens der Individuen treten andere Verzerrungen auf. In dem bereits erwähnten Sonderfall unverzerrter aber variabler Antworten der Befragten gilt $b_j = 0$ für $j = 1, 2, \dots, N$, d. h. $\bar{b} = 0$. Die Schätzung ist dann erwartungstreu.

In $\bar{b}$ können sich die individuellen Verzerrungen z. T. gegenseitig ausgleichen. Extreme b_j werden durch die Mittelwertbildung in $\bar{b}$ durch geringere bzw. gegengerichtete Verzerrungen geglättet. Selbst ein völliger Ausgleich der Fehler kann nicht ausgeschlossen werden. Diese Ausgleichstendenz unterscheidet sich jedoch von der Eigenschaft der Zufallsfehler, denn sie ist unabhängig vom Stichprobenumfang.

Der Stichprobenumfang bestimmt allerdings die zufällige Abweichung der in der Zufallsstichprobe tatsächlich realisierten Antwortfehler von der erwarteten Verzerrung $\bar{b}$. Diese zufällige Abweichung - entstanden durch den Auswahlprozeß in der Stichprobe - wird in der Varianz des Schätzers mit erfaßt. Es erscheint sinnvoll, diese Abweichung als Zufallsfehler zu be-

trachten, obwohl die Ursache, die individuelle Verzerrung, ein systematischer Fehler ist.

Der Zufallsfehler der Schätzung wird durch die Varianz des Schätzers gemessen. Neben der Stichprobenauswahl wirkt auch die Antwortvariabilität auf $\text{Var}(\bar{Y})$.

Mit (4.7) läßt sich die Varianz des Schätzers $\bar{Y}$ umformen, wobei wir

$$\frac{1}{n} \sum_i \varepsilon_i = \bar{\varepsilon} \quad (4.21)$$

setzen. In die Rechnung geht ein, daß die Antwortfehler ε_i unabhängig von den wahren Merkmalsausprägungen sind. Man erhält

$$\begin{aligned} \text{Var}(\bar{Y}) &= \text{Var}\left(\frac{1}{n} \sum_i Y_i\right) \\ &= \text{Var}\left(\frac{1}{n} \sum_i X_i + \frac{1}{n} \sum_i \varepsilon_i\right) \\ &= \text{Var}(\bar{X}) + \text{Var}(\bar{\varepsilon}). \end{aligned} \quad (4.22)$$

Nach (3.14) gilt für die Varianz des Schätzers der wahren Merkmalsausprägungen

$$\text{Var}(\bar{X}) = \frac{1}{n} \sigma_x^2 (1 - a^* f^*). \quad (4.23)$$

Zur besseren Unterscheidung der einzelnen Varianzen ist bei σ_x^2 zusätzlich der Index x eingefügt.

Um die Varianz des Stichprobenmittelwertes der Antwortfehler angeben zu können, bedarf es noch weiterer Überlegungen:

$$\begin{aligned} \text{Var}(\bar{\varepsilon}) &= \text{Var}\left(\frac{1}{n} \sum_i \varepsilon_i\right) \\ &= \frac{1}{n}, \text{Var}\left(\sum_i \varepsilon_i\right). \end{aligned}$$

Die Varianzaufspaltung, wobei als Bedingung eine gezogene Stichprobe $\underline{U}$ aus G zu unterstellen ist, führt zu

$$\text{Var}(\bar{E}) = \frac{1}{n} E[\text{Var}(\sum_i \epsilon_i | \underline{U})] + \frac{1}{n} \text{Var}[E(\sum_i \epsilon_i | \underline{U})]. \quad (4.24)$$

Innerhalb der eckigen Klammern wird dem Zufallsprozeß der Antwortvariabilität Rechnung getragen, während der Auswahlprozeß erst im zweiten Schritt Berücksichtigung findet.

Bei vorgegebener Stichprobe sind die ϵ_i jedoch unabhängige Zufallsvariablen, denn wir haben unterstellt, daß die Antworten der Individuen von denen anderer Individuen unabhängig sind. Auch wenn beim Ziehen mit Zurücklegen ein Element mehrmals in die Stichprobe gelangt, sollen die Antworten dieses Elements unabhängig voneinander sein¹⁾. Nur über den Auswahlprozeß entsteht eine Abhängigkeit. Deshalb gilt mit (4.12) und (4.3)

$$\begin{aligned} E[\text{Var}(\sum_i \epsilon_i | \underline{U})] &= E[\sum_i \text{Var}(\epsilon_i | \underline{U})] \\ &= E[\sum_i \text{Var}(\epsilon_i | U_i)] \\ &= \sum_i E[\text{Var}(\epsilon_i | U_i)] \\ &= \sum_i \sum_j \frac{1}{N} \sigma_j^2 \\ &= \sum_i \bar{\sigma}^2 \\ &= n \bar{\sigma}^2, \end{aligned} \quad (4.25)$$

Dieser Teil der Varianz von $\bar{Y}$ bleibt unbeeinflusst vom Auswahlprozeß. Anders sieht es bei der zweiten Komponente in (4.24) aus:

1) Auf diese Annahme wird im Anschluß an die Ableitung von $\text{Var}(\bar{Y})$ nochmals eingegangen.

$$\begin{aligned}\text{Var} \left[E \left(\sum_i \varepsilon_i \mid \underline{U} \right) \right] &= \text{Var} \left[\sum_i E(\varepsilon_i \mid \underline{U}) \right] \\ &= \text{Var} \left[\sum_i E(\varepsilon_i \mid U_i) \right].\end{aligned}$$

Die bedingten Erwartungswerte sind beim Auswahlvorgang Ziehen ohne Zurücklegen keine unabhängigen Zufallsvariablen; sie verhalten sich jedoch wie die X_i , denn sie sind wie diese an die Untersuchungseinheiten u_j gekoppelt. In Analogie zur einfachen Zufallsauswahl ohne Antwortfehler können wir deshalb mit (3.14) und (4.2) schreiben:

$$\begin{aligned}\text{Var} [E(\sum_i \varepsilon_i \mid \underline{U})] &= n (1-a f^*) \text{Var} [E(\varepsilon_i \mid U_i)] \\ &= n (1-a f^*) \frac{1}{N} \sum_j (b_j - \bar{b})^2 \\ &= n (1-a f^*) \sigma_b^2.\end{aligned}\quad (4.26)$$

Die Varianz von $\bar{\varepsilon}$ ergibt sich durch Einsetzen von (4.25) und (4.26) in (4.24):

$$\text{Var}(\bar{\varepsilon}) = \frac{1}{n} (1 - a f^*) \sigma_b^2 + \frac{1}{n} \bar{\sigma}^2. \quad (4.27)$$

In Verbindung mit (4.22) und (4.23) läßt sich damit die Varianz des Schätzers $\bar{Y}$ angeben:

$$\text{Var}(\bar{Y}) = \frac{1}{n} (1-f^* \cdot a) (\sigma_x^2 + \sigma_b^2) + \frac{1}{n} \bar{\sigma}^2. \quad (4.28)$$

Die Summe der Varianz σ_x^2 der Merkmalsausprägungen und der Varianz σ_b^2 der individuellen Verzerrungen kann man als Varianz der erwarteten Antworten interpretieren. Nur sie, nicht aber die durchschnittliche Varianz $\bar{\sigma}^2$ der Antworten, wird beim Ziehen ohne Zurücklegen durch die Endlichkeitskorrektur beeinflusst. Dies überrascht nicht, denn die Varianz der Antworten des befragten Elements wirkt nur innerhalb der Antworten eines Individuums, nicht aber im Auswahlprozeß der Stichprobe.

Ein Vergleich mit der Situation ohne Antwortfehler zeigt, daß die Varianz des Schätzers durch die Antwortfehler steigt:

$$\text{Var}(\bar{Y}) = \text{Var}(\bar{X}) + \frac{1}{n} (1 - af^*) \sigma_b^2 + \frac{1}{n} \bar{\sigma}^2.$$

Der Einfluß der Antwortfehler auf die Varianz des Schätzers $\bar{Y}$ hängt von den individuellen Verzerrungen und von der Variabilität der Antworten der u_j ab. Die bereits erwähnten Sonderfälle führen zu einer Reduktion der Einflußfaktoren.

(1) Verzerrungsfreie aber variable Antworten: Da $b_j = 0$, $j = 1, 2, \dots, N$, gilt, folgt $\sigma_b^2 = 0$ und damit

$$\text{Var}(\bar{Y}) = \text{Var}(\bar{X}) + \frac{1}{n} \bar{\sigma}^2.$$

Nur die Antwortvariabilität erhöht die Varianz des Schätzers.

(2) Verzerrte, aber konstante (nicht variable) Antworten: Die Bedingungen $b_j \neq 0$ und $\sigma_j^2 = 0$, $j = 1, 2, \dots, N$, führen zu

$$\text{Var}(\bar{Y}) = \text{Var}(\bar{X}) + \frac{1}{n} (1 - af^*) \bar{\sigma}_b^2.$$

Das Modell der einfachen Zufallsauswahl ohne Antwortfehler entsteht, wenn (1) und (2) kombiniert eintreten. Mit $b_j = 0$ und $\sigma_j^2 = 0$, $j = 1, 2, \dots, N$ folgt

$$\text{Var}(\bar{Y}) = \text{Var}(\bar{X}).$$

In die bisherigen Ausführungen geht ein, daß (beim Ziehen mit Zurücklegen) auch die Antworten von mehrmals in $\underline{U}$ enthaltenen Einheiten unabhängig voneinander sind. Es ist damit ein Erinnerungseffekt an die bereits durchgeführte Befragung ausgeschlossen. Diese Annahme entspricht sicherlich nicht ganz der Realität.

Das andere Extrem ist, wenn man voraussetzt, daß die Antworten mehrmals gezogener Elemente immer mit dem gleichen Fehler behaftet sind, d. h. sie erinnern sich an die bereits durchgeführte Befragung und antworten gleich¹⁾. Die Annahme beeinflußt nicht den Erwartungswert des Schätzers, nur $\text{Var}(\bar{Y})$. In die Beweisführung von Pokropp²⁾ geht diese Annahme ein, denn er gibt bei der Berechnung von $\text{Var}(\bar{Y})$ die Antworten der Elemente aus G als Bedingung vor³⁾. Das Ergebnis in der hier gewählten Schreibweise lautet:⁴⁾

$$\text{Var}(\bar{Y})_{mZ}^* = \frac{1}{n} (\sigma_x^2 + \sigma_h^2) + \frac{1}{n} \bar{\sigma}^2 (1 + \frac{n-1}{N}).$$

Da beim Ziehen ohne Zurücklegen die Elemente nicht mehrmals in die Stichprobe gefangen können, wirken sich die unterschiedlichen Annahmen nur beim Ziehen mit Zurücklegen (mZ) aus. Ein Vergleich mit (4.28) und $a = 0$ ergibt:

$$\text{Var}(\bar{Y})_{mZ}^* - \text{Var}(\bar{Y})_{mZ} = \frac{1}{n} \frac{n-1}{N} \bar{\sigma}^2.$$

Wenn sich die mehrmals befragten Einheiten noch an die vorherigen Antworten erinnern können, ist die Varianz des Schätzers größer, denn die Antworten dieser Einheiten sind dann positiv korreliert.

Die beiden hier behandelten Annahmen stellen nur zwei Möglichkeiten des Antwortverhaltens der in der Stichprobe mehrmals enthaltenen Elemente dar. Auch Mischformen zwischen beiden oder auch bewußte Veränderungen der Antworten sind denkbar. Der Einfluß auf die Varianz des Schätzers ist, zumindest bei großen Gesamtheiten, vernachlässigbar klein, so daß detailliertere Betrachtungen unterbleiben können.

1) Bzw. die Befragung wird aus Kostengründen nicht mehrmals durchgeführt, aber mehrmals ausgewertet.

2) Vgl.: Pokropp, 80, S. 207 ff.

3) Vgl.: Pokropp, 80, S. 208 ff.

4) Vgl.: Pokropp, 80, S. 209. Die unterschiedliche Annahme bezüglich des Antwortverhaltens der mehrmals in der Stichprobe enthaltenen Elemente beim Ziehen mit Zurücklegen (mZ) ist durch den Zusatz '*' hervorgehoben.

In der Literatur zu den Antwortfehlern in Stichproben wird teilweise auch eine Abhängigkeit zwischen den ε_j^i zugelassen. Diese Abhängigkeit gilt auch ohne den Auswahlprozeß der Stichprobe für die Elemente der Gesamtheit. Es gilt dann

$$\text{Cov}(\varepsilon_j^i, \varepsilon_k^i) \neq 0, \quad j, k = 1, 2, \dots, N \\ j \neq k$$

für mindestens ein Paar (u_j, u_k) der Elemente aus G. Mit

$$r = \frac{1}{N(N-1)} \sum_{\substack{j, k \\ j \neq k}} \frac{\text{Cov}(\varepsilon_j^i, \varepsilon_k^i)}{\bar{\sigma}^2}$$

ist dann der durchschnittliche Korrelationskoeffizient der Antwortfehler gegeben.

Die Abhängigkeit hat natürlich Einfluß auf die Varianz des Schätzers $\bar{Y}$. Ohne Beweis¹⁾ sei hier das Ergebnis angegeben:

$$\text{Var}(\bar{Y})_{mZ}^* = \frac{1}{n}(\sigma_x^2 + \sigma_b^2) + \frac{1}{n}\bar{\sigma}^2 \frac{N-1}{N} (1 + \frac{n}{N-1} + (n-1)r)$$

für Stichproben, die mit Zurücklegen gezogen werden, und

$$\text{Var}(\bar{Y})_{oZ}^* = \frac{1}{n}(1-f^*)(\sigma_x^2 + \sigma_b^2) + \frac{1}{n}\bar{\sigma}^2 (1 + (n-1)r)$$

für ohne Zurücklegen gezogene Stichproben.

Bei positiver Abhängigkeit zwischen den Fehlern, wie sie in den meisten Fällen auftritt²⁾, steigt die Varianz des Schätzers durch die Abhängigkeit an.

1) Vgl.: Pokropp, oC, S. 209 f; Cochran, 72, S. 444 f. Bei Pokropp wird beim Ziehen mit Zurücklegen für die mehrmals in die Stichprobe gelangenden Elemente vorausgesetzt, daß diese immer mit dem gleichen Antwortfehler antworten.
2) Vgl.: Cochran, 72, S. 445.

4.2.2.2 Die Abschätzbarkeit des Zufallsfehlers der Schätzung bei Antwortfehlern

Ein Maß für den Zufallsfehler einer Schätzung ist die Varianz des Schätzers. Liegen Antwortfehler vor, so erfaßt $\text{Var}(\bar{Y})$ neben dem durch die Stichprobe bedingten Zufallsfehler der wahren Merkmalsausprägungen auch die durch die Fehler in den Antworten bedingten zufälligen Abweichungen des Schätzers von seinem Erwartungswert.

Eine wichtige Eigenschaft des Zufallsfehlers ist seine Abschätzbarkeit aus der Stichprobe. Dies gilt jedenfalls für den Zufallsfehler der Schätzung, sofern keine Antwortfehler vorliegen, denn (3.16) gibt einen erwartungstreuen Schätzer für $\text{Var}(\bar{X})$ an. Zu überprüfen ist, ob die Abschätzbarkeit des Zufallsfehlers auch bei Antwortfehlern erhalten bleibt, d. h. ob in Analogie zu (3.16)

$$\widehat{\text{Var}}(\bar{Y}) = \frac{1}{n} (1 - af) S_Y^2 \quad (4.29)$$

$$\text{mit } S_Y^2 = \frac{1}{n-1} \sum_i (Y_i - \bar{Y})^2$$

ein erwartungstreuer Schätzer für $\text{Var}(\bar{Y})$ ist. Wegen

$$E(\widehat{\text{Var}}(\bar{Y})) = \frac{1}{n} (1 - af) E(S_Y^2) \quad (4.30)$$

müssen wir den Erwartungswert der korrigierten Stichprobenvarianz S_Y^2 berechnen. Die Zerlegung der Zufallsvariablen Y_i nach (4.7) ermöglicht eine Aufspaltung von S_Y^2 , so daß für $E(S_Y^2)$ gilt:

$$\begin{aligned} E(S_Y^2) &= E\left(\frac{1}{n-1} \sum_i (Y_i - \bar{Y})^2\right) \\ &= E\left(\frac{1}{n-1} \sum_i (X_i - \bar{X} + \varepsilon_i - \bar{\varepsilon})^2\right). \end{aligned}$$

Weil Unabhängigkeit zwischen Antwortfehlern und Merkmalsausprägung vorausgesetzt ist, fallen die gemischten Terme bei der Erwartungswertbildung weg. Es folgt daher:

$$E(S_Y^2) = E\left(\frac{1}{n-1} \sum_i (X_i - \bar{X})^2\right) + E\left(\frac{1}{n-1} \sum_i (\epsilon_i - \bar{\epsilon})^2\right).$$

Der erste Summand ist der Erwartungswert der korrigierten Stichprobenvarianz S^2 der wahren Merkmalsausprägungen (vgl. (3.15)); mit der analog definierten korrigierten Stichprobenvarianz S_ϵ^2 der Antwortfehler ϵ_i können wir kurz schreiben

$$E(S_Y^2) = E(S^2) + E(S_\epsilon^2). \quad (4.31)$$

Bei einfacher Zufallsauswahl ohne Antwortfehler ist gezeigt, daß

$$E(S^2) = \left(1 + \frac{a}{N-1}\right) \sigma_x^2$$

gilt. Es ist daher nur noch $E(S_\epsilon^2)$ zu berechnen, wobei die Annahme der Unabhängigkeit zwischen den Antwortfehlern ϵ_i , $i = 1, 2, \dots, n$, mit eingeht:

$$\begin{aligned} E(S_\epsilon^2) &= E\left(\frac{1}{n-1} \sum_i (\epsilon_i - \bar{\epsilon})^2\right) \\ &= \frac{1}{n-1} \left(\sum_i E(\epsilon_i^2) - 2E(\bar{\epsilon} \sum_i \epsilon_i) + \sum_i E(\bar{\epsilon}^2) \right) \\ &= \frac{1}{n-1} \left(\sum_i E(\epsilon_i^2) - 2nE(\bar{\epsilon}^2) + nE(\bar{\epsilon}^2) \right) \\ &= \frac{1}{n-1} \left(\sum_i E(\epsilon_i^2) - nE(\bar{\epsilon}^2) \right) \\ &= \frac{1}{n-1} \left(\sum_i \text{Var}(\epsilon_i) + \sum_i (E(\epsilon_i))^2 - n \text{Var}(\bar{\epsilon}) - n (E(\bar{\epsilon}))^2 \right). \end{aligned}$$

Für $E(\epsilon_i)$ und $E(\bar{\epsilon})$ ergibt sich mit (4.14) bzw. (4.20) das arithmetische Mittel $\bar{b}$ der individuellen Antwortverzerrungen b_j , so daß weiter folgt

$$E(S_\epsilon^2) = \frac{1}{n-1} \left(\sum_i \text{Var}(\epsilon_i) + n \bar{b}^2 - n \text{Var}(\bar{\epsilon}) - n \bar{b}^2 \right)$$

$$E(S_{\epsilon}^2) = \frac{1}{n-1} (\sum_i \text{Var}(\epsilon_i) - n \text{Var}(\bar{\epsilon})). \quad (4.32)$$

Mit (4.17), (4.27) und $f^* = \frac{n-1}{N-1}$ kann man weiter umformen:

$$\begin{aligned} E(S_{\epsilon}^2) &= \frac{1}{n-1} (\sum_i \bar{\sigma}^2 + \sum_i \sigma_b^2 - n \frac{1}{n} (1 - a f^*) \sigma_b^2 - n \frac{1}{n} \bar{\sigma}^2) \\ &= \frac{1}{n-1} (n \bar{\sigma}^2 - \bar{\sigma}^2 + n \sigma_b^2 - (1 - a f^*) \sigma_b^2) \\ &= \bar{\sigma}^2 + \sigma_b^2 (1 + \frac{a}{N-1}). \end{aligned} \quad (4.33)$$

Zusammenfassend ergibt sich für $E(S_Y^2)$

$$E(S_Y^2) = (1 + \frac{a}{N-1}) (\sigma_x^2 + \sigma_b^2) + \bar{\sigma}^2 \quad (4.34)$$

und eingesetzt in (4.30) gilt dann für den Erwartungswert von $\widehat{\text{Var}}(\bar{Y})$:

$$E(\widehat{\text{Var}}(\bar{Y})) = \frac{1}{n} (1 - fa) ((1 + \frac{a}{N-1}) (\sigma_x^2 + \sigma_b^2) + \bar{\sigma}^2).$$

Für den Fall Ziehen mit Zurücklegen ist $a = 0$ zu setzen und es gilt wegen (4.28)

$$\begin{aligned} E(\widehat{\text{Var}}(\bar{Y})_{mZ}) &= \frac{1}{n} (\sigma_x^2 + \sigma_b^2 + \bar{\sigma}^2) \\ &= \text{Var}(\bar{Y})_{mZ}, \end{aligned}$$

d. h. die Schätzung der Varianz des Schätzers ist erwartungstreu.

Beim Ziehen ohne Zurücklegen gilt mit $a = 1$

$$\begin{aligned} E(\widehat{\text{Var}}(\bar{Y})_{oZ}) &= \frac{1}{n} (1 - \frac{n}{N}) ((1 + \frac{1}{N-1}) (\sigma_x^2 + \sigma_b^2) + \bar{\sigma}^2) \\ &= \frac{1}{n} \frac{N-n}{N} \frac{N}{N-1} (\sigma_x^2 + \sigma_b^2) + \frac{1}{n} (1 - \frac{n}{N}) \bar{\sigma}^2 \end{aligned}$$

$$\begin{aligned} E(\widehat{\text{Var}(\bar{Y})}_{\text{OZ}}) &= \frac{1}{n} (1 - f^*) (\sigma_x^2 + \sigma_b^2) + \frac{1}{n} \bar{\sigma}^2 - \frac{1}{N} \bar{\sigma}^2 \\ &= \text{Var}(\bar{Y}) - \frac{1}{N} \bar{\sigma}^2. \end{aligned} \quad (4.35)$$

Beim Ziehen ohne Zurücklegen ist also die Schätzung der Varianz des Schätzers $\bar{Y}$ aus der Stichprobe negativ verzerrt¹⁾. Die Unterschätzung entsteht durch die Anwendung der Endlichkeitskorrektur in $\text{Var}(\bar{Y})$ auch auf die durchschnittliche Antwortvarianz $\bar{\sigma}^2$, obwohl diese nicht durch den Auswahlprozeß beeinflußt wird. Bei großen Gesamtheiten ist die Verzerrung jedoch vernachlässigbar klein.

Wenn unterstellt wird, daß beim Ziehen mit Zurücklegen die mehrmals in die Stichprobe gelangenden Elemente immer mit dem gleichen Fehler antworten, ist auch bei diesem Auswahlmodus $\text{Var}(\bar{Y})_{\text{mZ}}^*$ um $-\frac{1}{N} \bar{\sigma}^2$ verzerrt²⁾.

Die für die Erhebungspraxis sehr wichtige Eigenschaft der Zufallsfehler, die Abschätzbarkeit aus der Stichprobe, bleibt auch bei Antwortfehlern, abgesehen von der geringfügigen Verzerrung beim Ziehen ohne Zurücklegen, erhalten. Mit $\text{Var}(\bar{Y})$ sind also nicht nur die durch die Stichprobenauswahl bedingten Zufallsfehler erfaßt, sondern auch die durch die Fehler in den Antworten.

4.2.3 Auswirkungen der Antwortfehler auf die Schätzung bei einfacher Zufallsauswahl

Zusammenfassend bleibt festzuhalten, daß die systematischen Fehler in den Antworten auf Grund der indi-

1) Vgl.: Pokropp, 80, S. 211; Cochran, 72, S. 443.

2) Vgl.: Pokropp, 80, S. 211.

viduellen Antwortverzerrungen b_j zu einem bias in der Schätzung führen bzw. führen können. Die Auswirkungen einer solchen Verzerrung auf die Schätzaussage, das Konfidenzintervall, wurden bereits in Abschnitt 3.2.4.1 eingehend beschrieben. Die Ergebnisse lassen sich auf die Antwortfehler übertragen, denn bezogen auf die Konfidenzaussage ist es prinzipiell gleichgültig, welche Ursache die Verzerrung der Schätzung hat.

Natürlich ist bei der Formulierung der Konfidenzaussage bei Antwortfehlern $\text{Var}(\bar{Y})$ statt $\text{Var}(\bar{X})$ einzusetzen, denn die relative Verzerrung, bezogen auf die Standardabweichung des Schätzers, ist für die Veränderung der Wahrscheinlichkeit der Konfidenzaussage durch die systematischen Fehler ausschlaggebend. In Analogie zu (3.30) können wir formulieren: Die tatsächliche Wahrscheinlichkeit, mit der das Konfidenzintervall den zu schätzenden Parameter μ überdeckt, sinkt durch die Antwortfehler mit zunehmendem Verhältnis zwischen Verzerrung $\bar{b}$ und der Standardabweichung des Schätzers $\bar{Y}$, wie es in Tabelle 3.2 angegeben ist.

Im Unterschied zur Situation mit konstanten systematischen Fehlern ist auch die Varianz des Schätzers durch die Fehler in den Antworten vergrößert, denn in $\text{Var}(\bar{Y})$ (vgl. (4.28)), ist sowohl der Einfluß der Antwortvariabilität über die durchschnittliche Antwortvarianz $\bar{\sigma}^2$ als auch ein Einfluß der individuellen Verzerrungen b_j über die Varianz σ_b^2 enthalten. Die b_j verzerren daher nicht nur die Schätzung, sie wirken sich auch negativ auf die Präzision der Schätzung aus.

Eine große Varianz der individuellen Verzerrungen muß nicht unbedingt mit einer starken Verzerrung der Schätzung gekoppelt sein. Sind einige b_j positiv

während andere Einheiten eine negative Antwortverzerrung aufweisen, führt das zu einer großen Varianz σ_b^2 . In dieser Situation heben sich die individuellen Verzerrungen in $\bar{b}$ zum Teil gegenseitig auf, sogar ein völliger Verzerrungsausgleich ist denkbar. Umgekehrt bedeutet eine große Verzerrung $\bar{b}$ nicht unbedingt auch ein großes σ_b^2 , denn im Extremfall, wenn alle b_j gleich sind, ist $\sigma_b^2 = 0$, unabhängig vom Ausmaß der Verzerrung der Schätzung.

Die Vergrößerung der Varianz des Schätzers durch die Antwortfehler beeinflusst auch die Auswirkungen der Verzerrung $\bar{b}$ auf die Schätzaussage, denn die relative Verzerrung sinkt mit steigender Varianz des Schätzers. Diesem (Schein)-Vorteil steht jedoch ein entscheidender Nachteil gegenüber: Die Grenzen des Konfidenzintervalls werden durch die Antwortfehler nach außen verschoben, das Konfidenzintervall wird länger.

Der Vergleich zwischen verzerrter und unverzerrter Schätzung über den Stichprobenumfang, wie er im Abschnitt 3.2.4.2 vorgenommen wurde, berücksichtigt die bei Antwortfehlern als Verzerrungsursache zusätzlich bestehende Verlängerung des Konfidenzintervalls nicht. Ein längeres Intervall kann bei einer Schätzung ohne Antwortfehler - gleiches Konfidenzniveau vorausgesetzt - bei einem geringeren Stichprobenumfang erreicht werden. Dies läßt sich durch folgende Überlegung zeigen.

Die Varianz des Schätzers bei Antwortfehlern kann man mit (4.22) umformen zu

$$\text{Var}(\bar{Y}) = \text{Var}(\bar{X}) \cdot C^2,$$

wobei

$$C^2 = \left[1 + \frac{\text{Var}(\bar{\epsilon})}{\text{Var}(\bar{X})} \right] \geq 1$$

gilt. Das Konfidenzintervall

$$\left[\bar{Y} \pm z_{\alpha/2} \cdot \sqrt{\text{Var}(\bar{Y})} \right] = \left[\bar{Y} \pm z_{\alpha/2} \cdot \sqrt{\text{Var}(\bar{X})} \cdot C \right]$$

ist um den Faktor C länger als ein entsprechendes Intervall ohne Antwortfehler. Für den Auswahlmodus 'Ziehen mit Zurücklegen' gilt mit (3.10) speziell

$$\left[\bar{Y} \pm z_{\alpha/2} \cdot \frac{C}{\sqrt{n}} \sigma \right].$$

In Analogie zu (3.32) läßt sich der Stichprobenumfang n'' berechnen, mit der eine vergleichbare Konfidenzaussage (gleiche tatsächliche Wahrscheinlichkeit, gleiche Intervalllänge) erreicht werden kann, wenn die Antwortfehler vermieden werden können ¹⁾

$$z_{\alpha/2} \cdot \frac{C}{\sqrt{n}} = \frac{z_{\alpha'/2}}{\sqrt{n''}}$$

$$\frac{n''}{n} = \left(\frac{z_{\alpha'/2}}{z_{\alpha/2}} \right)^2 \cdot \frac{1}{C^2}.$$

Wegen $C \geq 1$ ist $\frac{n''}{n}$ kleiner als ein entsprechendes Verhältnis ohne Verlängerung des Konfidenzintervalls durch die Antwortfehler. Die Werte der Tabelle 3.3 bzw. 3.4 sind daher mit dem Faktor $\frac{1}{C^2}$ zu multiplizieren, wenn neben der Verzerrung des Schätzers auch eine Vergrößerung der Varianz des Schätzers eintritt.

4.2.4 Die Eigenschaften des Zufallsfehlers der Schätzung bei Antwortfehlern

In Abschnitt 2.3.1 dieser Arbeit sind die Eigenschaften des Zufallsfehlers einer Schätzung ohne Antwort-

1) Für den Vergleich, bezogen auf eine vorgegebene Wahrscheinlichkeit gilt entsprechendes.

fehler beschrieben. Es ist zu überprüfen, inwieweit diese Eigenschaften (Abschätzbarkeit, Ausgleichstendenz, kein Zufallsfehler bei einer Totalerhebung) auch für den Zufallsfehler

$$Z_{\bar{Y}} = \bar{Y} - E(\bar{Y}),$$

der die Zufallskomponente in den Antworten der Untersuchungseinheiten mit erfaßt, gültig bleiben.

Für die Erhebungspraxis ist besonders wichtig, daß der Zufallsfehler der Schätzung auch bei Fehlern in den Antworten aus der Stichprobe geschätzt werden kann. Die Schätzfunktion (4.29) für $\text{Var}(\bar{Y}) = \text{Var}(Z_{\bar{Y}})$ weist zwar beim Ziehen ohne Zurücklegen¹⁾ eine geringfügige Verzerrung auf, die aber bei großen Gesamtheiten vernachlässigt werden kann.

Der Unterschätzung von $\text{Var}(\bar{Y})$ kann man durch Vernachlässigung der Endlichkeitskorrektur entgegenwirken, denn dadurch steigt $E(\widehat{\text{Var}(\bar{Y})})$. Es tritt dann allerdings eine Überschätzung der Varianz des Schätzers ein²⁾.

Die Ausgleichstendenz des Zufallsfehlers kommt durch den Faktor $\frac{1}{n}$ in der Varianzformel zum Ausdruck, d. h. auch bei Antwortfehlern läßt sich der Zufallsfehler der Schätzung durch eine Steigerung des Stichprobenumfangs reduzieren.

Der Grenzfall $n=N$ führt beim Ziehen ohne Zurücklegen zur Totalerhebung. Die Varianz des Schätzers ohne Antwortfehler sinkt dann auf Null, d. h. bei einer Totalerhebung besteht kein Zufallsfehler. Gilt dies auch bei Antwortfehlern?

1) Wenn man keine Unabhängigkeit zwischen den Antworten der mehrmals in die Stichprobe gelangten Untersuchungseinheiten unterstellen kann, gilt das auch für den Auswahlmodus 'Ziehen mit Zurücklegen'.

2) Vgl.: Cochran, 72, S. 443.

Setzen wir $n = N$ in die Formel (4.28) für die Varianz der Schätzung $\bar{Y}$ ein, so folgt für den Fall Ziehen ohne Zurücklegen:

$$\begin{aligned}\text{Var}(\bar{Y}) &= \frac{1}{N} \left(1 - \frac{N-1}{N-1}\right) (\sigma_x^2 + \sigma_b^2) + \frac{1}{N} \bar{\sigma}^2 \\ &= \frac{1}{N} \bar{\sigma}^2.\end{aligned}$$

Auch bei einer Totalerhebung ist es daher nicht möglich, die Varianz auf Null zu senken, d. h. die Totalerhebung ist bei Antwortfehlern eine Schätzung. Wenn die Totalerhebung als Schätzung noch einen Zufallsfehler aufweist, verliert sie gegenüber der Stichprobe weiter an Attraktivität, insbesondere auch deshalb, weil bei realistischen Annahmen mit einer Stichprobe vom Umfang $n = 0,25 N$ fast die gleiche Präzision wie bei einer Totalerhebung erreicht werden kann¹⁾. Zu beachten ist, daß in praktischen Untersuchungen die individuellen Verzerrungen b_j bzw. Antwortvarianzen σ_j^2 nicht unabhängig vom Stichprobenumfang sind, denn mit steigendem n kann nicht mehr die gleiche Sorgfalt auf die Befragung des einzelnen Individuums gelegt werden - die Folge: $|b_j|$ und/oder σ_j^2 nehmen zu. Nur mit steigenden Kosten ist eine Beibehaltung der Erhebungsqualität bei wachsendem Stichprobenumfang zu erreichen²⁾.

Unter der Voraussetzung, daß σ_b^2 bzw. $\bar{\sigma}^2$ mit n steigen, führt die Erhöhung des Stichprobenumfanges zwar weiterhin zu einer Reduktion des stichprobenbedingten Zufallsfehlers, der Einfluß der Antwortvariabilität wird jedoch größer. Insgesamt kann bei großen n die Varianz des Schätzers wieder ansteigen³⁾. Erschwerend kommt hinzu, daß die Totalerhebung unter diesen Voraussetzungen eine größere Verzerrung aufweist bzw. aufweisen kann.

1) Vgl.: Stenger, 77, S. 400 ff.

2) Vgl.: Stenger, 77, S. 399 ff; Pokropp, 80, S. 216 f.

3) Vgl.: Pokropp, 80, S. 216 f.

4.3 Antwortfehler im Schichtenauswahlverfahren

4.3.1 Grundlegende Notationen

Die Einflüsse der systematischen Fehler in den Daten auf die Ergebnisse einer geschichteten Stichprobe lassen sich aus den bisherigen Aussagen ableiten. Dabei ist das Modell der geschichteten Auswahl, wie es in Abschnitt 3.3 dargestellt ist, zugrundegelegt.

Die Elemente der Gesamtheit G sind in m Schichten G_h , $h = 1, 2, \dots, m$, elementfremd eingeteilt. Aus jeder Schicht wird eine Stichprobe nach dem Modell der einfachen Zufallsauswahl - unabhängig von den anderen Schichten - gezogen. Im Gegensatz zu Abschnitt 3.3 sind hier Antwortfehler unterstellt, d. h. die Befragung der ausgewählten Elemente ergibt nicht deren wahre Merkmalsausprägung sondern eine, mit einem zufälligen Antwortfehler überlagerte Ausprägung. In Analogie zu (2.12) läßt sich die Antwort der Untersuchungseinheit u_{hj} aus Schicht h durch die Zufallsvariable ¹⁾

$$Y'_{hj} = x_{hj} + \epsilon'_{hj}, \quad j = 1, 2, \dots, N_h, \quad (4.36)$$

$$h = 1, 2, \dots, m,$$

beschreiben.

Die individuelle Antwortverzerrung b_{hj} ergibt sich als Erwartungswert der individuellen Antwortfehler ϵ'_{hj}

$$E(\epsilon'_{hj}) = b_{hj}, \quad (4.37)$$

so daß für die erwartete Antwort gilt

1) Sukhatme u. Sukhatme übertragen ihr Modell mit Interviewereinfluß auch auf das Schichtenauswahlverfahren; vgl.: Sukhatme u. Sukhatme, 70, S. 402 f.

$$E(Y'_{hj}) = x_{hj} + b_{hj}. \quad (4.38)$$

Es existiert ebenfalls eine Varianz der individuellen Antworten

$$\text{Var}(Y'_{hj}) = \text{Var}(\epsilon'_{hj}) = \sigma_{hj}^2. \quad (4.39)$$

In Abschnitt 3.3 dieser Arbeit sind die zur Beschreibung des Schichtenauswahlverfahrens notwendigen Formeln und Symbole, bezogen auf die wahren Merkmalsausprägungen x_{hj} , bereits eingeführt. Diese gelten auch in diesem Abschnitt, wobei allerdings zur besseren Unterscheidung zwischen den Varianzen, die sich auf die individuellen Verzerrungen beziehen, und den Varianzen der x_{hj} letztere zusätzlich durch einen Index 'x' gekennzeichnet sind. Daher bezeichnet (vgl. (3.45))

$$\sigma_{xh}^2 = \sigma_h^2 = \frac{1}{N_h} \sum_h (x_{hj} - \mu_h)^2$$

die Varianz der wahren Merkmalsausprägungen in Schicht h und (vgl. (3.47))

$$\sigma_x^2 = \sigma^2 = \frac{1}{N} \sum_h \sum_j (x_{hj} - \mu)^2 =$$

$$\sum_h W_h \sigma_{xh}^2 + \sum_h W_h (\mu_h - \mu)^2$$

ist die Varianz der wahren Merkmalsausprägungen der Einheiten in der Gesamtheit.

In der folgenden Übersicht sind die zur Berücksichtigung der Antwortfehler im Schichtenauswahlverfahren notwendigen Formeln und Symbole zusammengestellt:

Durchschnittliche individuelle Verzerrung in Schicht h

$$\bar{b}_h = \frac{1}{N_h} \sum_j b_{hj}, \quad h = 1, 2, \dots, m; \quad (4.40)$$

durchschnittliche Verzerrung in der Gesamtheit

$$\bar{b}_g = \sum_n w_h \bar{b}_h = \frac{1}{N} \sum_h \sum_j b_{hj}; \quad (4.41)$$

Varianz der individuellen Verzerrungen in Schicht h

$$\sigma_{bh}^2 = \frac{1}{N_h} \sum_j (b_{hj} - \bar{b}_h)^2, \quad h = 1, 2, \dots, m; \quad (4.42)$$

Varianz der Verzerrung zwischen den Schichten

$$\sigma_{bz}^2 = \sum_h w_h (\bar{b}_h - \bar{b}_g)^2; \quad (4.43)$$

Varianz der Verzerrungen in G

$$\sigma_b^2 = \frac{1}{N} \sum_h \sum_j (b_{hj} - \bar{b}_g)^2 = \sigma_{bw}^2 + \sigma_{bz}^2; \quad (4.44)$$

durchschnittliche Antwortvarianz in Schicht h

$$\bar{\sigma}_h^2 = \frac{1}{N_h} \sum_j \sigma_{hj}^2; \quad (4.45)$$

durchschnittliche Antwortvarianz in G

$$\bar{\sigma}^2 = \sum_h w_h \bar{\sigma}_h^2 = \frac{1}{N} \sum_h \sum_j \sigma_{hj}^2. \quad (4.46)$$

4.3.2 Die geschichtete Mittelwertschätzung bei Antwortfehlern

4.3.2.1 Erwartungswert und Varianz des geschichteten Mittelwertschätzers bei Antwortfehlern

Da in den einzelnen Schichten unabhängig voneinander die Stichproben nach dem Modell der einfachen Zufallsauswahl gezogen werden, lassen sich die Ergebnisse

aus Abschnitt 4.2, zur Unterscheidung der einzelnen Schichten zusätzlich mit dem Index h versehen, übernehmen. Die Stichprobe $\underline{U}_h$ vom Umfang n_h aus G_h führt zu dem Zufallsvektor

$$\begin{aligned}\underline{Y}_h &= (Y_{hi}/i = 1, 2, \dots, n_h) \\ &= (X_{hi}/i = 1, 2, \dots, n_h) + (\epsilon_{hi}/i = 1, 2, \dots, n_h)\end{aligned}\quad (4.47)$$

mit

$$Y_{hi} = Y'_{hj},$$

falls u_{hj} beim Auswahlvorgang i in Schicht h gezogen wurde. Aus $\underline{Y}_h$ ergibt sich der Schätzer für den Mittelwert μ_h der wahren Merkmalsausprägungen in Schicht h als

$$\begin{aligned}\bar{Y}_h &= \frac{1}{n_h} \sum_i Y_{hi} \\ &= \frac{1}{n_h} \sum_i X_{hi} + \frac{1}{n_h} \sum_i \epsilon_{hi} \\ &= \bar{X}_h + \bar{\epsilon}_h.\end{aligned}\quad (4.48)$$

Die Schätzung ist verzerrt, denn mit (4.38) gilt in Analogie zu (4.20)

$$E(\bar{Y}_h) = \mu_h + \bar{b}_h. \quad (4.49)$$

Die Bildung des Erwartungswertes erfolgt über die möglichen Antworten des ausgewählten Elements und anschließend über den Auswahlvorgang.

Die Varianz des Mittelwertschätzers $\bar{Y}_h$ in Schicht h ist wegen (4.28)

$$\text{Var}(\bar{Y}_h) = \frac{1}{n_h} (1-f_h^* a_h) (\sigma_{xh}^2 + \sigma_{bh}^2) + \frac{1}{n_h} \bar{\sigma}_h^2. \quad (4.50)$$

Sie kann aus der Stichprobe mit (vgl. (4.29))

$$\widehat{\text{Var}}(\bar{Y}_h) = \frac{1}{n_h} (1-f_h a_h) S_{Yh}^2 \quad (4.51)$$

geschätzt werden, wobei

$$S_{Yh}^2 = \frac{1}{n_h - 1} \sum_i (Y_{hi} - \bar{Y}_h)^2$$

die korrigierte Stichprobenvarianz aus Schicht h bedeutet. Dabei ist zu beachten, daß beim Ziehen ohne Zurücklegen eine geringfügige Unterschätzung von $\text{Var}(\bar{Y}_h)$ besteht¹⁾.

Ein Schätzer $\bar{Y}_g$ für den Parameter μ der Gesamtheit G ergibt sich in Analogie zu (3.54) als das mit den relativen Schichtenumfängen W_h gewogenes arithmetisches Mittel der Schätzer $\bar{Y}_h$:

$$\bar{Y}_g = \sum_h W_h \bar{Y}_h. \quad (4.52)$$

Auch beim geschichteten Schätzer für μ führen die individuellen Verzerrungen zu einem bias, denn mit (4.49), (3.44) und (4.41) folgt

$$\begin{aligned} E(\bar{Y}_g) &= \sum_h W_h E(\bar{Y}_h) \\ &= \sum_h W_h (\mu_h + \bar{b}_h) \\ &= \mu + \bar{b}_g. \end{aligned} \quad (4.53)$$

Dabei ist die Verzerrung $\bar{b}_g$ das gewogene arithmetische Mittel der Schichtenverzerrungen. Sie ist identisch

1) Vgl. Abschnitt 4.2.2.2 dieser Arbeit.

mit dem Durchschnitt aller individuellen Verzerrungen in G. Gegengerichtete individuelle Verzerrungen wirken daher ausgleichend auf die Gesamtverzerrung¹⁾. Theoretisch ist sogar ein völliger Ausgleich denkbar.

Die Varianz des geschichteten Schätzers bei Antwortfehlern läßt sich mit (4.52) und (4.50) angeben:

$$\begin{aligned}\text{Var}(\bar{Y}_g) &= \sum_h W_h^2 \text{Var}(\bar{Y}_h) \\ &= \sum_h W_h^2 \left(\frac{1}{n_h} (1 - f_h a_h) (\sigma_{xh}^2 + \sigma_{bh}^2) + \frac{1}{n_h} \bar{\sigma}_h^2 \right).\end{aligned}\quad (4.54)$$

Analog zur Situation ohne Antwortfehler beeinflussen nur die Varianzen der Antworten innerhalb der Schichten (über σ_{xh}^2 , σ_{bh}^2 und $\bar{\sigma}_h^2$) die Varianz des Schätzers $\bar{Y}_g$ beim Schichtenauswahlverfahren. Weil aus jeder Schicht eine Teilschätzung für den jeweiligen Schichtenmittelwert erfolgt, haben die Unterschiede zwischen den Schichten, gemessen durch die Varianz der Schichtenmittelwerte μ_h bzw. Schichtenverzerrungen $\bar{b}_h$ zwischen den Schichten, keinen Einfluß auf $\text{Var}(\bar{Y}_g)$.

Für die Erhebungspraxis ist die Abschätzbarkeit der Varianz aus der Stichprobe von großer Bedeutung. Mit (4.51) ist

$$\begin{aligned}\widehat{\text{Var}}(\bar{Y}_g) &= \sum_h W_h^2 \widehat{\text{Var}}(\bar{Y}_h) \\ &= \sum_h W_h^2 \cdot \frac{1}{n_h} (1 - f_h a_h) S_{Yh}^2\end{aligned}\quad (4.55)$$

ein Schätzer für $\text{Var}(\bar{Y}_g)$, wobei die Unterschätzung der Schichtenvarianzen und damit auch der Varianz des geschichteten Schätzers beim Auswahlmodus 'Ziehen ohne Zurücklegen' in den Schichten zu beachten ist²⁾.

-
- 1) Positive und negative Schichtenverzerrungen führen ebenfalls zu einer Reduzierung bzw. unter Umständen zu einem Ausgleich der Verzerrung in $\bar{b}_g$. Dieser Effekt läßt sich jedoch auf die individuellen Verzerrungen zurückführen, denn diese bestimmen über (4.41) sowohl $\bar{b}_h$ als auch $\bar{\sigma}_h^2$.
 - 2) Zu den Bedingungen, unter denen die Schätzung beim Ziehen mit Zurücklegen erwartungstreu ist, vgl. Abschnitt

4.3.2.2 Die Aufteilung der Stichprobe auf die Schichten unter Berücksichtigung der Antwortfehler

Die Varianz des geschichteten Mittelwertschätzers ist über die Stichprobenumfänge n_h abhängig von der Aufteilung der Gesamtstichprobe auf die einzelnen Schichten¹⁾. Das gilt auch für die Schichtenauswahl bei Antwortfehlern. Die beiden besonders wichtigen Aufteilungskriterien, 'größenproportionale' und 'optimale' Aufteilung, sind bereits in Abschnitt 3.3.1.2 erwähnt worden. Es ist zu untersuchen, inwieweit die Fehler in den Antworten der Untersuchungseinheiten die Aufteilung beeinflussen.

Bei größenproportionaler Aufteilung sind ausschließlich die relativen Schichtenumfänge bestimmend für die n_h . Die Antwortfehler haben daher keinen Einfluß auf diese Art der Aufteilung der Stichprobe. Anders sieht es bei optimaler Aufteilung aus, denn die Antwortfehler beeinflussen die Varianz des geschichteten Schätzers. Für den Auswahlmodus 'Ziehen mit Zurücklegen' lassen sich die optimalen Stichprobenumfänge in den Schichten durch²⁾

$$n_{h,opt} = \frac{w_h \sigma_{xh}}{\sum_h w_h \sigma_{xh}} n$$

berechnen, sofern keine Antwortfehler vorliegen. Die Varianz des geschichteten Schätzers $\bar{Y}_g$ ohne Antwortfehler (vgl. (3.55))

$$V_{\text{opt}}(\bar{Y}_g)_{\text{opt}} = \sum_h w_h^2 \frac{1}{n_h} \sigma_{xh}^2$$

ist dann minimal, nicht aber die Varianz des Schätzers $\bar{Y}_g$

- 1) Der Einfluß der Schichteneinteilung auf die Varianz des Schätzers bleibt hier unberücksichtigt.
- 2) Bei großen Schichtenumfängen, wenn $N_h - 1$ in Näherung gleich N_h gesetzt werden kann, gilt die Formel auch für den Auswahlmodus 'Ziehen ohne Zurücklegen'. Zur optimalen Aufteilung vgl. z. B. Menges u. Skala, 73, S. 105 f.

bei Antwortfehlern. Um $\text{Var}(\bar{Y}_g)$ zu minimieren, müssen die Antwortfehler mit berücksichtigt werden. Setzt man

$$\sigma_{Yh}^2 = \sigma_{xh}^2 + \sigma_{bh}^2 + \bar{\sigma}_h^2 = \text{Var}(Y_{hi}),$$

so gilt mit (4.54) für $a_h = 0$, $h = 1, 2, \dots, m$,

$$\text{Var}(\bar{Y}_g) = \sum_h W_h^2 \frac{1}{n_h} \sigma_{Yh}^2.$$

In Analogie zur Situation ohne Antwortfehler folgt dann

$$n_{h, \text{opt}, Af} = n \frac{\sigma_{Yh} W_h}{\sum_h W_h \sigma_{Yh}}.$$

Bei der optimalen Aufteilung der Stichprobe auf die Schichten ist daher nicht nur die Varianz der wahren Merkmalsausprägungen innerhalb der Schichten zu beachten, sondern auch die individuellen Verzerrungen und die durchschnittlichen Antwortvarianzen innerhalb der Schichten.

4.3.3 Auswirkungen der Antwortfehler auf die Schätzungsaussage beim Schichtenauswahlverfahren

Die Auswirkungen der systematischen Fehler in den Antworten der Untersuchungseinheiten lassen sich in Analogie zu Abschnitt 4.2.3 beschreiben. Zwei Aspekte sind zu beachten: die Wirkung des durch die individuellen Antwortverzerrungen verursachten Bias der Schätzung auf die Schätzungsaussage und der Einfluß der Antwortfehler auf die Präzision der Schätzung.

Die Verteilung des Schätzers $\bar{Y}_g$ ist wegen der Antwortfehler um den Bias $\bar{b}_g$ von dem zu schätzenden Parameter μ weg verschoben. Die Konfidenzaussage

$$\bar{Y}_g - z_{\alpha/2} \sqrt{\text{Var}(\bar{Y}_g)} \leq \mu \leq \bar{Y}_g + z_{\alpha/2} \sqrt{\text{Var}(\bar{Y}_g)}$$

tritt daher nicht mehr mit der vorgegebenen Wahrscheinlichkeit $\gamma = 1 - \alpha$, sondern mit $\gamma'(\bar{b}_g) < \gamma$ ein. Diese tatsächliche Wahrscheinlichkeit der Intervallaussage sinkt mit zunehmender relativer Verzerrung

$$b_{g,Af}^* = \frac{\bar{b}_g}{\sqrt{\text{Var}(\bar{Y}_g)}}$$

gemäß Tabelle 3.2, wobei darin $b_{g,Af}^*$ für b^* zu setzen ist.

Ein Vergleich zwischen der Varianz des geschichteten Schätzers mit und ohne Antwortfehler zeigt, daß durch die individuellen Antwortfehler eine Vergrößerung der Varianzen des Schätzers $\bar{Y}_g$ erfolgt. Durch Einsetzen von (3.55) in (4.54) ergibt sich

$$\text{Var}(\bar{Y}_g) = \text{Var}(\bar{X}_g) + \sum_h w_h^2 \frac{1}{n_h} ((1-f_h^* a_h) \sigma_{bh}^2 + \bar{\sigma}_h^2). \quad (4.56)$$

Die Auswirkungen der Antwortfehler auf die Varianz sind vor allem dann groß, wenn die individuellen Verzerrungen in den Schichten stark voneinander abweichen (σ_{bh}^2 groß) und wenn die durchschnittliche Antwortvarianz $\bar{\sigma}_h^2$ in den Schichten groß ist. Unterschiede in den Schichtenverzerrungen $\bar{b}_h$, gemessen durch die Varianz σ_{bz}^2 (vgl. (4.43)) haben keinen Einfluß auf die Varianz des Schätzers $\bar{Y}_g$.

Die Vergrößerung der Varianz des geschichteten Schätzers führt zu einer Verlängerung des Konfidenzintervalls, denn mit steigender Varianz verschieben sich die Intervallgrenzen nach außen. Bei einer Gegenüberstellung des Stichprobenumfanges n einer geschichteten Schätzung mit Antwortfehlern und dem Stichprobenumfang n' , der ausgereicht hätte, eine vergleichbare Schätz-

aussage (gleiche Intervalllänge, gleiche tatsächliche Wahrscheinlichkeit) zu erreichen, wenn keine Antwortfehler vorliegen, muß die Steigerung des Zufallsfehlers durch die Fehler in den Antworten berücksichtigt werden.

In Analogie zu den Ausführungen in Abschnitt 4.2.3 ist durch die Varianzvergrößerung wegen der Antwortfehler mit einer stärkeren Reduzierung des notwendigen Stichprobenumfanges n' auch beim Schichtenauswahlverfahren zu rechnen. Auf eine formale Ableitung soll jedoch verzichtet werden.

Die Ausführungen zu den Eigenschaften des Zufallsfehlers der Schätzung mit Antwortfehlern bei einfacher Zufallsauswahl (Abschnitt 4.2.4) lassen sich auf den geschichteten Schätzer übertragen. Eine detaillierte Beschreibung kann daher unterbleiben.

4.4 Vergleich zwischen einfacher Zufallsauswahl und Schichtenauswahlverfahren unter Berücksichtigung systematischer Fehler in den Antworten

4.4.1 Annahmen, die dem Vergleich zugrunde liegen

Neben dem bisher betrachteten Einfluß der Antwortfehler auf die Stichprobenverfahren ist auch eine Gegenüberstellung zwischen einfacher Zufallsauswahl und geschichteter Auswahl, jeweils mit Fehlern in den Antworten, von Interesse. Es geht darum zu überprüfen, ob die Vorteile des Schichtenauswahlverfahrens auch bei Antwortfehlern gültig bleiben, oder ob sie sich durch die systematischen Fehler noch verstärken oder abschwächen. Der Vergleich der Verfahren soll ebenso wie in Abschnitt 3.4 über den Schichteneffekt erfolgen. Das hat den Vorteil, daß durch Einbeziehung des Schichteneffekts ohne Antwortfehler auch die Einflüsse dieser Fehler auf die Verbesserung der Schätzung durch die Schichtung deutlich werden.

Um die Unterschiede in den Ergebnissen beider Verfahren auf die systematischen Fehler in den Antworten zurückführen zu können, sind die Einflüsse anderer Parameter auf die Ergebnisse auszuschließen. Andernfalls ist ein Vergleich nicht sinnvoll, da die Unterschiede dann nicht nur durch die Verfahren selbst, sondern durch die unterschiedlichen Annahmen bei den Verfahren entstehen. Im einzelnen sei unterstellt:

- (1) gleicher Stichprobenumfang bei beiden Verfahren, d. h.

$$\sum_h n_h = n;$$

- (2) gleiche individuelle Verzerrung der einzelnen Untersuchungseinheiten im jeweiligen Verfahren, d. h.

$$b_j = b_{ht} \text{ falls } u_j = u_{ht}$$

$$j = 1, 2, \dots, N$$

$$t = 1, 2, \dots, N_h; h = 1, 2, \dots, m;$$

- (3) gleiche Varianz der individuellen Antworten der Untersuchungseinheiten

$$\sigma_j^2 = \sigma_{ht}^2 \text{ falls } u_j = u_{ht}$$

Natürlich muß vorausgesetzt werden, daß jeweils die gleiche Gesamtheit betrachtet wird und beim Vergleich der Schichteneffekte mit und ohne Antwortfehler die Schichtung gleich ist.

Die Annahme (1) wurde bereits in Abschnitt 3.4 unterstellt und die Annahmen (2) und (3) erscheinen auf den ersten Blick ebenfalls unproblematisch, denn bei gleicher Gesamtheit entspricht jedem u_j aus G genau ein u_{ht} aus $\bigcup_h G_h$. Die Untersuchungseinheiten haben, so wird unterstellt, ein konstantes Antwortverhalten mit der Verzerrung b_j und der Varianz σ_j^2 , unabhängig vom Stichprobenverfahren.

Implizit ist damit unterstellt, daß bei beiden Verfahren das gleiche Erhebungssystem benutzt wird, daß z. B. bei Interviewerbefragungen die gleichen Interviewer bzw. solche mit gleichem Ausbildungsstand die gleichen Fragen stellen. Gerade diese Annahme wird bei praktischen Untersuchungen oft nicht erfüllt sein, denn der Vorteil des Schichtenauswahlverfahrens ist es, eine bestimmte Schätzgenauigkeit bei geringerem Stichprobenumfang zu erreichen. Dies wiederum ermög-

licht ein Erhebungssystem mit besser geschulten Interviewern, mit differenzierteren Fragen, mit umfangreicherem Erhebungsziel.

Wegen des oft geringeren Stichprobenumfanges beim Schichtenauswahlverfahren sind auch weitergehende Maßnahmen zur Vermeidung systematischer Fehler, so auch der Antwortfehler, möglich, ebenso wie Maßnahmen zur Bestimmung der Verzerrung.

Der praktische Vergleich zweier Erhebungen¹⁾ ist daher fast unmöglich. Theoretisch kann jedoch durch die oben genannten Annahmen überprüft werden, bei welchem Verfahren die Gefahren durch die Antwortfehler größer sind.

Aus den Annahmen (2) und (3) lassen sich Folgerungen ableiten: Es gilt wegen (2)

$$\bar{b} = \frac{1}{N} \sum_j b_j = \frac{1}{N} \sum_h \sum_t b_{ht} = \bar{b}_g, \quad (4.57)$$

d. h. die Gesamtverzerrung der Schätzung ist bei beiden Verfahren gleich. Analog läßt sich aus (3) ableiten

$$\bar{\sigma}^2 = \frac{1}{N} \sum_j \sigma_j^2 = \frac{1}{N} \sum_h \sum_t \sigma_{ht}^2 = \sum_h w_h \bar{\sigma}_h^2. \quad (4.58)$$

Diese Beziehungen sind wichtig zur Umformung des Schichteneffektes.

4.4.2 Die Beziehung zwischen der Varianz in der Gesamtheit und der Varianz innerhalb der Schichten

Für den Vergleich zwischen einfacher Zufallsauswahl und geschichteter Auswahl im Schichteneffekt ist eine Beziehung zwischen der Varianz der Zufallsvariablen

1) Wenn die Erhebungen nicht speziell zur vergleichenden Analyse der systematischen Fehler in Stichprobenverfahren durchgeführt werden.

Y_i (Auswahl aus der Gesamtheit) und der Varianz der Zufallsvariablen Y_{hi} (Auswahl innerhalb der Schichten) notwendig. In Abschnitt 3.4 erfüllt die Aufspaltung der Varianz der Merkmalsausprägungen (vgl. (3.47)) in der Gesamtheit

$$\begin{aligned}\sigma^2 = \text{Var}(X_i) &= \sum_h W_h \sigma_{xh}^2 + \sum_h W_h (\mu_h - \mu)^2 \quad (4.59) \\ &= \sum_h W_h \sigma_{xh}^2 + \sigma_{xz}^2\end{aligned}$$

diese Funktion. Wegen

$$\text{Var}(Y_i) = \text{Var}(X_i) + \text{Var}(\varepsilon_i) \quad (4.60)$$

ist noch eine vergleichbare Aufspaltung für $\text{Var}(\varepsilon_i)$ abzuleiten.

In der Zufallsvariablen ε_i kommt der Auswahlvorgang unter den Elementen aus G und deren Antwortverhalten zum Ausdruck. Bei der Zufallsauswahl der u_j aus G ist es ein Zufallsereignis, aus welcher Schicht das gezogene u_j stammt; die Schichtung stellt eine Zerlegung dar. Geben wir eine Schicht h als Bedingung bei der Auswahl vor, so können wir $\text{Var}(\varepsilon_i)$ schreiben als¹⁾

$$\text{Var}(\varepsilon_i) = E[\text{Var}(\varepsilon_i | G_h)] + \text{Var}[E(\varepsilon_i | G_h)]. \quad (4.61)$$

Mit

$$E(\varepsilon_i | G_h) = \bar{b}_h$$

folgt weiter

$$\begin{aligned}\text{Var}[E(\varepsilon_i | G_h)] &= \sum_h W_h (\bar{b}_h - \bar{b})^2 \\ &= \sigma_{bz}^2.\end{aligned} \quad (4.62)$$

1) Zur Varianzaufspaltung vgl. z. B. Pokropp, 80, S. 233 f.

Die Bildung der Varianz erfolgt über die Bedingungen, d. h. die Schichtung, wobei die Wahrscheinlichkeit für das Eintreten der Bedingungen durch die W_h gegeben sind.

Den ersten Term aus (4.61) können wir ähnlich umformen, wobei (4.42) und (4.45) zur Anwendung kommen:

$$\begin{aligned}\text{Var}(\epsilon_i | G_h) &= \text{Var}(\epsilon_{ht}) \\ &= \sigma_{bh}^2 + \bar{\sigma}_h^2.\end{aligned}$$

Bei der Berechnung von $\text{Var}(\epsilon_{ht})$ muß sowohl das Antwortverhalten der ausgewählten Einheiten, als auch die Auswahl der Stichprobe berücksichtigt werden.

Die Erwartungswertbildung über die Bedingungen führt in Verbindung mit (4.46) zu

$$\begin{aligned}E[\text{Var}(\epsilon_i | G_h)] &= \sum_h W_h (\sigma_{bh}^2 + \bar{\sigma}_h^2) \\ &= \sum_h W_h \sigma_{bh}^2 + \bar{\sigma}^2, \quad (4.63)\end{aligned}$$

Gleichung (4.63) gibt die durchschnittliche Varianz der Antwortfehler innerhalb der Schichten an. Hier wirken sowohl die individuellen Verzerrungen als auch die Antwortvariabilität. Im Gegensatz dazu ist die Varianz der Antwortfehler zwischen den Schichten nur von den Schichtenverzerrungen beeinflusst.

Durch Einsetzen der Gleichungen (4.62) und (4.63) in (4.61) folgt

$$\text{Var}(\epsilon_i) = \sum_h W_h (\sigma_{hb}^2 + \bar{\sigma}_h^2) + \sigma_{bz}^2. \quad (4.64)$$

In Verbindung mit (4.59) ergibt sich wegen (4.60)

$$\text{Var}(Y_i) = \sum_h W_h (\sigma_{xh}^2 + \sigma_{bh}^2 + \bar{\sigma}_h^2) + \sigma_{xz}^2 + \sigma_{bz}^2. \quad (4.65)$$

4.4.3 Der Schichteneffekt unter Berücksichtigung der Antwortfehler

Der Schichteneffekt D ist nach (3.62) definiert als Differenz zwischen der Varianz des Schätzers bei einfacher Zufallsauswahl und bei geschichteter Auswahl. Für die Mittelwertschätzung bei Antwortfehlern folgt deshalb

$$D_{Af} = \text{Var}(\bar{Y}) - \text{Var}(\bar{Y}_g) \quad (4.66)$$

und mit (4.28) und (4.54), wobei wir uns auf den Fall 'Ziehen mit Zurücklegen' beschränken¹⁾,

$$D_{Af} = \frac{1}{n}(\sigma_x^2 + \sigma_b^2 + \bar{\sigma}^2) - \sum_h W_h^2 \frac{1}{n_h}(\sigma_{xh}^2 + \sigma_{bh}^2 + \bar{\sigma}_h^2).$$

Die Varianzaufspaltung nach (4.65) ermöglicht in Verbindung mit (4.18) die Umformung

$$\begin{aligned} D_{Af} &= \frac{1}{n} \sum_h W_h (\sigma_{xh}^2 + \sigma_{bh}^2 + \bar{\sigma}_h^2) + \frac{1}{n}(\sigma_{xz}^2 + \sigma_{bz}^2) \\ &\quad - \sum_h W_h^2 \frac{1}{n_h} (\sigma_{xh}^2 + \sigma_{bh}^2 + \bar{\sigma}_h^2) \\ &= \frac{1}{n}(\sigma_{xz}^2 + \sigma_{bz}^2) + \frac{1}{Nn} \sum_h W_h (\sigma_{xh}^2 + \sigma_{bh}^2 + \bar{\sigma}_h^2) \left(\frac{N}{n} - \frac{N_h}{n_h} \right). \quad (4.67) \end{aligned}$$

Unter Beachtung von (3.64) kann man auch schreiben:

$$D_{Af} = D_x + \frac{1}{n} \sigma_{bz}^2 + \frac{1}{N} \sum_h W_h (\sigma_{bh}^2 + \bar{\sigma}_h^2) \left(\frac{N}{n} - \frac{N_h}{n_h} \right).$$

Der Schichteneffekt D_x der wahren Merkmalsausprägungen,

1) Die Ergebnisse können auch auf das Ziehen ohne Zurücklegen übertragen werden, wenn die Auswahlssätze f_h in den Schichten klein sind (z. B. $f_h < 0,05$).

er entspricht dem Schichteneffekt D ohne systematische Fehler nach (3.64), ist durch die Antwortfehler um zwei Summanden vergrößert. Diese lassen sich als Schichteneffekt D_{ε} der Antwortfehler bezeichnen. Mit

$$D_{\varepsilon} = \frac{1}{n} \sigma_{bz}^2 + \frac{1}{N} \sum_h W_h (\sigma_{bh}^2 + \bar{\sigma}_h^2) \left(\frac{N}{n} - \frac{N_h}{n_h} \right) \quad (4.68)$$

kann man daher auch schreiben

$$D_{Af} = D_x + D_{\varepsilon}.$$

Beim Schichteneffekt mit Antwortfehlern ist hervorzuheben, daß nur in dem, von der Aufteilung der Stichprobe auf die Schichten abhängigen, zweiten Term in (4.67) die durchschnittliche Antwortvarianz $\bar{\sigma}_h^2$ Einfluß hat.

In der für die Praxis besonders wichtigen größenproportionalen Aufteilung der Stichprobe auf die Schichten ist der Schichteneffekt wegen

$$\frac{N}{n} = \frac{N_h}{n_h}$$

auf

$$D_{Af, \text{prop}} = \frac{1}{n} (\sigma_{xz}^2 + \sigma_{bz}^2) = D_x + \frac{1}{n} \sigma_{bz}^2$$

reduziert. Der Schichteneffekt ist durch die Antwortfehler um $\frac{1}{n} \sigma_{bz}^2$ gestiegen.

Aus (4.67) wird nochmals deutlich, daß bei der Bestimmung der optimalen Aufteilung neben der Varianz der wahren Merkmalsausprägungen auch die Antwortfehler zu berücksichtigen sind.

Sind die Varianzen innerhalb der Schichten stark unterschiedlich, so empfiehlt sich die optimale Aufteilung, weil dadurch die Varianz des Schätzers $\bar{Y}$ ver-

ringert werden kann. Der Schichteneffekt ist dann größer als $D_{Af,prop}$, denn die Unterschiede zwischen den σ_{bh}^2 gehen in den Schichteneffekt ein.

Der bereits mehrmals angesprochene Sonderfall unverzerrter, aber variabler Antworten führt wegen $b_j = 0$ für $j = 1, 2, \dots, N$ zu einer Reduzierung des Schichteneffekts der Antwortfehler. Wegen $\sigma_{bh}^2 = \sigma_{bz}^2 = 0$ folgt aus (4.68) speziell:

$$D_E = \frac{1}{N} \sum_h W_h \bar{\sigma}_h^2 \left(\frac{N}{n} - \frac{N_h}{n_h} \right).$$

Nur die durchschnittliche Varianz der individuellen Antworten in den Schichten bestimmt die Veränderung des Schichteneffektes durch die Antwortfehler. Diese ist abhängig von der Aufteilung der Stichprobe auf die Schichten; bei proportionaler Aufteilung gilt:

$$D_{E,prop} = 0 \text{ bzw. } D_{Af,prop} = D_{x,prop}.$$

In diesem durchaus realistischen Spezialfall bleibt der Schichteneffekt demnach von der Antwortvariabilität unberührt.

Insgesamt bleibt festzuhalten, daß die Vorteile des Schichtenauswahlverfahrens auch bei systematischen Fehlern in den Daten erhalten bleiben. Allerdings muß bei der Interpretation der Ergebnisse aus Teilgesamtheiten die Verzerrung der Schätzung der Schichtenmittelwerte beachtet werden.

Besonders wichtig ist, daß auch bei Antwortfehlern das Schichtenauswahlverfahren eine, im Vergleich zur einfachen Zufallsauswahl, bessere Schätzung liefert; der Schichteneffekt ist bei proportionaler Aufteilung und günstiger Schichtung positiv. Die Antwortfehler können

sogar den Schichteneffekt vergrößern ($D_c > 0$), d. h. das Schichtenauswahlverfahren nutzt die Antwortfehler zu einer Verbesserung der Schätzung aus. Damit wird der Varianzvergrößerung durch die systematischen Fehler in den Antworten entgegengewirkt, eine Aufhebung ist jedoch nicht möglich.

Zusätzlich ist zu beachten, daß ein positiver Schichteneffekt wiederum die negative Wirkung der Verzerrung auf die Schätzaussage verstärkt. Dies wurde ja bereits in Abschnitt 3.4.2.2 für die Verzerrung des Schätzers $\bar{X}_g$ ohne Antwortfehler gezeigt. Besteht bei einfacher und geschichteter Zufallsauswahl jeweils die gleiche absolute Verzerrung $\bar{b} = \bar{b}_g$, so führt der positive Schichteneffekt beim Schichtenauswahlverfahren zu einer höheren relativen Verzerrung. Dieses Verhältnis zwischen Verzerrung und Standardabweichung des Schätzers ist nach (3.30) entscheidend für die Auswirkung der Verzerrung auf die Schätzaussage.

Die Ergebnisse der Berechnungen aus Abschnitt 3.4.2.2 lassen sich auch auf den Vergleich zwischen einfacher Zufallsauswahl und Schichtenauswahl bei Antwortfehlern übertragen (vgl. Tabelle 3.6, Schaubilder 3.4 und 3.5). Auf eine detailliertere Betrachtung kann daher verzichtet werden.

5. Schlußbetrachtungen

Zielsetzung dieser Arbeit war es, Fehler und Fehlerquellen, die bei der Durchführung einer Stichprobenerhebung auftreten können, und die Auswirkung von Fehlern auf die Schätzaussage zu untersuchen.

Die Ausführungen in Kapitel 2 geben einen Überblick über die möglichen Fehler in Stichprobenerhebungen. Dabei wurden zwischen Zufallsfehlern, systematischen Fehlern und der Verzerrung der Schätzung unterschieden. Die systematischen Fehler, d. h. die Fehler, die auf Mängel und Unzulänglichkeiten in der Durchführung der Erhebung zurückzuführen sind, erzeugen eine Verzerrung des Schätzers. Einige systematische Fehler, die mit den Antworten der Untersuchungseinheiten entstehen, beeinflussen auch den Zufallsfehler. An Hand der Eigenschaften der Fehlerarten wurden die Unterschiede zwischen ihnen nochmals herausgestellt.

In Kapitel 3 und 4 dieser Arbeit wurden die Auswirkungen verschiedener Fehlerursachen auf eine Schätzung, speziell der Mittelwertschätzung, bei zwei Stichprobenverfahren - der einfachen Zufallsauswahl und der Schichtenauswahl - behandelt.

Die Vielzahl der verschiedenen Fehler in Stichprobenerhebungen machte eine Beschränkung auf einzelne Fehlerarten notwendig. Näher untersucht wurden zum einen Fehler, die bei korrekten Antworten der Befragten auftreten (Kapitel 3) und zum anderen Fehler in den Antworten der Untersuchungseinheiten (Kapitel 4). In beiden Fällen wurde unterstellt, daß alle ausgewählten Untersuchungseinheiten die gestellte Frage beantworten, d. h. das Phänomen der Nichtbeantwortung blieb in den Behandlungen unberücksichtigt.

Hier sollen noch einmal zusammenfassend die Auswirkungen der aus den Fehlerursachen entstehenden Verzerrungen auf die Schätzaussage der Stichprobenerhebung dargestellt werden.

Die Schätzaussage einer Stichprobenerhebung erfolgt in der Regel als Wahrscheinlichkeitsaussage bezogen auf ein Intervall, in dem der zu schätzende Parameter liegt - das Konfidenzintervall. Die Lage des Intervalls in der Menge der Ausprägungsmöglichkeiten wird durch den Schätzer bzw. seinen Erwartungswert bestimmt; das Konfidenzniveau, der Verteilungstyp und die Varianz des Schätzers legen die Intervalllänge fest.

Ist die Schätzung durch die vielfältigen Fehlerursachen verzerrt, so ist das Konfidenzintervall, bezogen auf den zu schätzenden Parameter, verschoben. Damit sinkt die tatsächliche Wahrscheinlichkeit, mit der das Konfidenzintervall den wahren Parameter überdeckt. Die Verzerrung b ist hier als Gesamtverzerrung zu interpretieren, als Summe der Verzerrungen verschiedener Ursachen, z. B. individuelle Antwortfehler und Verzerrung der Auswahlgrundlage. Unter Umständen können sich die Biaskomponenten z. T. gegenseitig aufheben, wenn die verschiedenen Ursachen gegengerichtete Verzerrungen erzeugen. Entscheidend für die Verringerung der Wahrscheinlichkeit des Konfidenzintervalls ist das Verhältnis zwischen Gesamtverzerrung und der Standardabweichung des Schätzers, die relative Verzerrung b^* .

Ist die Verteilung des Schätzers symmetrisch (z. B. Normalverteilung), so genügt es, den Betrag des Bias zu betrachten, denn positive und negative Verzerrungen haben, bezogen auf die Wahrscheinlichkeit des Konfidenzintervalls, die gleiche Auswirkung. In Kapitel

3 ist die Auswirkung der Verzerrung auf einen normalverteilten Schätzer für den Mittelwert in der Gesamtheit in Tabelle 3.2 zusammengefaßt. Dabei wurde die einfache Zufallsauswahl zugrundegelegt.

Die Tabelle ist jedoch auch für andere normalverteilte Schätzer gültig, auch wenn sie mit anderen Stichprobenverfahren ermittelt sind. Der Ansatz des Konfidenzintervalls ist unabhängig vom zu schätzenden Parameter und vom Stichprobenverfahren.

Für eine Verzerrung, die die Hälfte der Standardabweichung des Schätzers beträgt ($b^* = 0,5$), gibt die Tabelle bei einem Konfidenzniveau von 0,95 eine Reduzierung der Wahrscheinlichkeit von rund 3 % auf 0,92091 an. Der Rückgang der Wahrscheinlichkeit steigt mit zunehmender Verzerrung überproportional: Beträgt das Verhältnis $b^* = 1$, so sinkt die Wahrscheinlichkeit bei gleichem Konfidenzniveau bereits um mehr als 12 % auf 0,82993. Geringfügige relative Verzerrungen (z. B. $b^* < 0,25$) beeinflussen die Schätzaussage so wenig, daß eine Vernachlässigung gerechtfertigt erscheint.

Die Betrachtungen der Auswirkungen systematischer Fehler auf das Konfidenzintervall bilden den Ausgangspunkt einer Gegenüberstellung zwischen einer verzerrten und einer unverzerrten Mittelwertschätzung. Es wurde untersucht, welcher Stichprobenumfang ausgereicht hätte, eine vergleichbare Schätzaussage zu erreichen, wenn die systematischen Fehler vermieden werden könnten. Eine vergleichbare Schätzaussage bedeutet in diesem Zusammenhang, daß bei verzerrter und unverzerrter Schätzung jeweils ein Konfidenzintervall gleicher Länge entsteht, das auch mit der gleichen tatsächlichen Wahrscheinlichkeit den zu schätzenden Parameter μ überdeckt.

In Tabelle 3.3 und Tabelle 3.4 sind die Ergebnisse eines solchen Vergleichs¹⁾, einmal bezogen auf die tatsächliche Wahrscheinlichkeit der verzerrten Intervallschätzung und zum anderen auf das vorgegebene Konfidenzniveau, angegeben. Gehen wir von einer Erhebung aus, deren Ergebnis eine relative Verzerrung von $b^* = 1$ aufweist. Durch entsprechende Verlängerung des Konfidenzintervalls überdeckt dieses den zu schätzenden Parameter μ mit der vorgegebenen Wahrscheinlichkeit $\gamma = 0,95$. Könnte die Verzerrung vermieden werden, so hätte eine Erhebung mit rund 55% des Stichprobenumfangs der verzerrten Schätzung ausgereicht, ein Konfidenzintervall gleicher Länge zu erzeugen.

Beim Schichtenauswahlverfahren ist zur Umrechnung der Auswirkungen der Verzerrung in den Vergleich der Stichprobenumfänge eine Annahme über die Aufteilung der Gesamtstichprobe auf die Schichten notwendig. Bei proportionaler und optimaler Aufteilung ist eine entsprechende Umrechnung möglich, so daß die Tabellen 3.3 und 3.4 auch für das Schichtenauswahlverfahren anwendbar sind.

Das Schichtenauswahlverfahren ermöglicht im Vergleich zur einfachen Zufallsauswahl eine genauere Schätzung, d. h. die Varianz des geschichteten Schätzers ist geringer als die des Schätzers bei einfacher Zufallsauswahl bei gleichem Stichprobenumfang. Ist bei beiden Verfahren der Schätzer jeweils um den gleichen Betrag verzerrt, so führt das zu einer höheren relativen Verzerrung des Schätzers beim Schichtenauswahlverfahren. In Abschnitt 3.4 wurde untersucht, wie stark dieser Effekt die Schätzaussage bei der geschichteten Auswahl beeinflußt.

Tabelle 3.6 gibt die tatsächliche Wahrscheinlichkeit, mit der das Konfidenzintervall um den verzerrten Schätzer den zu schätzenden Parameter μ überdeckt, in Abhängigkeit von der relativen Verzerrung b^* und dem relativen Schichteneffekt D^* an.

1) Die Berechnungen sind auf die einfache Zufallsauswahl mit Zurücklegen ausgerichtet.

Besteht beispielsweise eine relative Verzerrung von $b^* = 0,5$, so beträgt die tatsächliche Wahrscheinlichkeit des Konfidenzintervalls ($\gamma = 0,95$) noch 0,92091. Kann die Varianz der Schätzung durch die Schichtung um 40 % reduziert werden ($D^* = 0,4$), so überdeckt das Intervall um den Schätzer den Parameter μ noch mit einer tatsächlichen Wahrscheinlichkeit von rund 0,9. Für die Erhebungspraxis ist wichtig, daß auch beim Schichtenauswahlverfahren geringfügige relative Verzerrungen (z. B. $b^* < 0,25$) vernachlässigt werden können, solange der relative Schichteneffekt D^* unter 0,8 bleibt.

Ist die Schätzung durch Antwortfehler verzerrt, so ist zusätzlich zur Verzerrung auch die Vergrößerung des Zufallsfehlers zu beachten. Vergleicht man das Ergebnis einer verzerrten mit dem einer unverzerrten Stichprobe über den Stichprobenumfang, so ist dieser Effekt zu berücksichtigen. Durch die Vermeidung der Antwortfehler ist daher eine zusätzliche Reduzierung des Stichprobenumfangs möglich.

Für die Erhebungspraxis ist von besonderer Bedeutung, daß trotz der Antwortfehler die Varianz des Schätzers aus der Stichprobe geschätzt werden kann. Zwar besteht beim Auswahlmodus 'Ziehen ohne Zurücklegen' eine negative Verzerrung ($-\frac{1}{N} \bar{O}$), die aber bei großen Gesamtheiten vernachlässigt werden kann.

Der Zufallsfehler, der durch die Fehler in den Antworten entsteht, führt dazu, daß auch bei einer Totalerhebung noch ein Zufallsfehler besteht - die Totalerhebung wird zu einer Schätzung.

In dieser Arbeit konnten nicht alle Aspekte, die im Zusammenhang mit systematischen Fehlern in Stichproben von Bedeutung sind, angesprochen werden. Insbesondere blieben die Möglichkeiten und Grenzen der

Vermeidung von systematischen Fehlern bzw. die Ermittlung der resultierenden Verzerrung unberücksichtigt. Insgesamt wurde jedoch deutlich, daß den Mängeln und Unzulänglichkeiten in einer Stichprobenerhebung große Aufmerksamkeit geschenkt werden muß.

LEBENS LAUF

Angaben zur Person

Geburtsdatum	23. April 1953
Geburtsort	Gindorf, Kreis Bitburg/Prüm
Staatsangehörigkeit	deutsch
Familienstand	ledig

Schulbildung

Schuleintritt 1959	Volksschule Gindorf
1959 bis 1964	Volksschule Trier-Heiligkreuz
1964 bis 1969	Staatliche Realschule Trier
1969 bis 1972	Wirtschaftsgymnasium Trier

Grundwehrdienst

Juli 1972 bis September 1973

Studium

WS 73/74 bis SS 79	Wirtschaftspädagogik an der Johannes Gutenberg-Universität Mainz
Mai 1979	Diplomprüfung

vom 1. 10. 1979 bis 30. 9. 1984	wissenschaftlicher Mitarbeiter am Fachbereich Rechts- und Wirtschaftswissenschaften der Johannes Gutenberg-Universität Mainz
------------------------------------	--

19. 12. 1984	mündliche Doktorprüfung
--------------	-------------------------

Mainz, den 10. 1. 1985

Johes Stroud

DRUCKEREI GRÄBNER
GOTENSTRASSE 3
8602 ALTENDORF
Tel. 09545 / 87 91

